Issues in Target Tracking

Peter Willett
ECE Department
University of Connecticut, Storrs, CT 06269, USA
willett@engr.uconn.edu

Abstract

In this lecture we discuss a number of concerns and items of interest that related to the tracking of targets in clutter. We begin with a discussion of performance evaluation methods. The familiar CRLB is shown to be adjusted in a straightforward way when deterministic trajectories are estimated in the presence of measurement-origin uncertainty (false alarms and missed detections); indeed, the same is true for nondeterministic ones, meaning those with process noise. We then mention the HYCA (hybrid conditional averaging) approach, which has considerable appeal in that it can predict the track-life as opposed to the accuracy of tracks that are kept. We then discuss track testing: the sequential probability ratio test (SPRT) for track acceptance, the Page test for track deletion, and here we most especially discuss variants (such as Shiryaev) for the case of fluctuating targets, as would be found in multi-static (fused) systems. Finally, we mention some management issues, specifically aspects of sampling time for multi-sensor systems, sensor placement with target tracking in mind, and the choice of waveform when the goal is tracking.

I. Performance Evaluation

A. The Static Case: The CRLB When Measurements are of Uncertain Origin

1) Introduction: In many estimation situations measurements are of uncertain origin. This is best exemplified by the target-tracking situation in which at each scan a number m_t of measurements are obtained, and it is not known which, if any, of these is target-originated. In several earlier papers the surprising observation was made that the Cramér-Rao lower bound (CRLB) for the estimation of a fixed parameter vector (e.g., initial position and velocity) that characterizes the target motion, for the special case multidimensional measurements in the presence of additive white Gaussian noise, is simply a multiple of that for the case with no uncertainty. That is, there is a scalar *information-reduction factor*. Details of this material are available in [24].

2) The Multi-Parameter CRLB: In many estimation problems one is faced with the problem that one's data is of uncertain origin. For example, in the target-tracking situation, the data set may consist of "hits" (threshold exceedances) indicative of a target's presence in a particular location at a particular time; however, hits may be of spurious origin (that is, they are *false-alarms*), and indeed it may be that the true target is unrepresented in the data set at the time in question (a *missed detection*). At issue is how well one can estimate a fixed parameter vector that characterizes target motion given such uncertainty.

The measure in which we are interested is, not unnaturally, the mean-squared error (MSE) of the estimate. There is a classical result for this, known as the Cramér-Rao lower bound (CRLB) (e.g. [20], [36]). Let us assume access to an observation $\mathbf{Z}$ which has probability density function (pdf) $p(\mathbf{Z};\mathbf{x})$, meaning that the pdf depends on a parameter vector $\mathbf{x}$ which is to be estimated. Let us assume the existence of an *unbiased* estimator $\hat{\mathbf{x}}(\mathbf{Z})$, meaning that we have

$$\mathcal{E}\left\{\hat{\mathbf{x}}(\mathbf{Z})\right\} = \mathbf{x} \tag{1}$$

in which $\mathcal{E}$ denotes expectation parametrized by $\mathbf{x}$. Then under fairly broad regularity conditions the CRLB has it that

$$\mathcal{E}\left\{\left[\hat{\mathbf{x}}(\mathbf{Z})-\mathbf{x}\right]\left[\hat{\mathbf{x}}(\mathbf{Z})-\mathbf{x}\right]^T\right\} \geq \mathbf{J}^{-1} \tag{2}$$

in which

$$\mathbf{J} \equiv \mathcal{E}\left\{\left[\nabla_{\mathbf{x}}\log\left(p(\mathbf{Z};\mathbf{x})\right)\right]\left[\nabla_{\mathbf{x}}\log\left(p(\mathbf{Z};\mathbf{x})\right)\right]^T\right\} \tag{3}$$

is Fisher's information matrix. Again under broad regularity conditions, if a maximum-likelihood estimator (MLE) for $\mathbf{x}$ exists, then it achieves the CRLB asymptotically.

3) Measurements of Uncertain Origin: The general multi-parameter CRLB is fairly standard; let us now turn to the case of particular interest. We define the aggregate observation

$$\mathbf{Z} = \{\mathbf{Z}(1),\mathbf{Z}(2),\ldots,\mathbf{Z}(T)\} \tag{4}$$

in which the tth observation is

$$\mathbf{Z}(t) = \{\mathbf{z}_i(t)\}_{i=1}^{m_t} \tag{5}$$

meaning, in the target-tracking situation, that there are m_t individual observations which comprise it. Although m_t is of course known to the estimator, to compute the MSE it is necessary to average over the possible values of m_t, and hence we assume

the existence of a probability mass function $q(m_t)$ controlling m_t, and of a related measure $\epsilon(m_t)$ denoting the probability that a particular measurement is target-originated, given that this true measurement is not missed. We further assume that

$$p(\mathbf{Z};\mathbf{x}) = \prod_{t=1}^{T} p(\mathbf{Z}(t);\mathbf{x}) \tag{6}$$

$$= \prod_{t=1}^{T}\left(\sum_{m_t=0}^{\infty} q(m_t)\left[\prod_{i=1}^{m_t} p_0(\mathbf{z}_i(t))\right]\left[(1-\epsilon(m_t)) + \frac{\epsilon(m_t)}{m_t}\sum_{i=1}^{m_t}\frac{p_1(\mathbf{z}_i(t)-\mu_t(\mathbf{x}))}{p_0(\mathbf{z}_i(t))}\right]\right) \tag{7}$$

which requires some explanation. First, we have assumed that, conditioned on $\mathbf{x}$, $\{\mathbf{Z}(1),\mathbf{Z}(2),\ldots,\mathbf{Z}(T)\}$ are independent. Second, we have assumed that, given there are m_t observations which comprise $\mathbf{Z}(t)$, any of m_t+1 events are possible: either all m_t observations $\{\mathbf{z}_i(t)\}_{i=1}^{m_t}$ are distributed according to $p_0(\cdot)$ (which does not depend on $\mathbf{x}$); or *exactly one* of these is distributed according to $p_1(\cdot)$ (which *does* depend on $\mathbf{x}$), while the rest remain distributed according to $p_0(\cdot)$, with each of these m_t events equally likely. Regardless of which event is true, all m_t observations are independent given that event. Third, we have written the dependence of the target-generated observations on the unknown parameter $\mathbf{x}$ as $\mu_t(\mathbf{x})$; while variations are possible, this model is most appropriate for a deterministic track (such as straight-line or ballistic motion) in which the target's location is a function of a few "initial" parameters such as position and velocity. Finally, it is important that the dependence of the observation $\mathbf{z}_i(t)$ on $\mu_t(\mathbf{x})$ is as a direct translation (mean-shift).

For example, in the target tracking scenario $\mathbf{Z}(t)$ is comprised of all observations collected at time t, and these observations can be all false-alarms (the detection from the target has been missed), or can contain exactly one true detection and (m_t-1) false alarms. In this situation it is common to assume $p_0(\mathbf{x})$ is uniform over the observation volume (or gated volume); and that the number of false-alarms is Poisson, meaning that

$$q(m_t) = (1-P_d)\frac{(\lambda V)^{m_t}e^{-\lambda V}}{m_t!} + P_d\frac{(\lambda V)^{(m_t-1)}e^{-\lambda V}}{(m_t-1)!} \tag{8}$$

$$\epsilon(m_t) = \frac{P_d}{q(m_t)}\frac{(\lambda V)^{(m_t-1)}e^{-\lambda V}}{(m_t-1)!} \tag{9}$$

are, respectively, the a-priori probability that there are m_t $(\in\{0,\infty\})$ observations at time t, and the probability that $\mathbf{z}_i(t)$ is target-generated given that there are m_t measurements at time t. In the above P_d is the probability of detection, λ is the average number of false-alarms per unit observation volume, and V is the actual observation volume.

4) The Information-Reduction Factor: In order to apply (7) to (2), we first note that due to independence of $\mathbf{Z}(t)$ and the logarithm we have

$$\mathbf{J} = \sum_{t=1}^{T}\mathbf{J}_t \tag{10}$$

in which

$$\mathbf{J}_t \equiv \mathcal{E}\left\{(\nabla_{\mathbf{x}}\log[p(\mathbf{Z}(t);\mathbf{x})])(\nabla_{\mathbf{x}}\log[p(\mathbf{Z}(t);\mathbf{x})])^T\right\} \tag{11}$$

Further, since the number of observations m_t is known to the estimator, we also have

$$\mathbf{J}_t = \sum_{m_t=1}^{\infty} q(m_t)\mathbf{J}_t(m_t) \tag{12}$$

in which

$$\mathbf{J}_t(m_t) = \mathcal{E}\left\{(\nabla_{\mathbf{x}}\log[p(\{\mathbf{z}_i(t)\}_{i=1}^{m_t};\mathbf{x})])(\nabla_{\mathbf{x}}\log[p(\{\mathbf{z}_i(t)\}_{i=1}^{m_t};\mathbf{x})])^T\right\} \tag{13}$$

reflects estimation efficiency at a particular time t for a particular number of observations m_t.

In [17], [18] a surprising result was obtained. Under the target-tracking assumptions above and further assuming a Gaussian model for the true-observation pdf $p_1(\cdot)$, then it is possible to write

$$\mathbf{J}_t(m_t) = q_2(P_d,\lambda V)\mathbf{J}_t^0 \tag{14}$$

where

$$\mathbf{J}_t^0 \equiv \mathcal{E}\left\{(\nabla_{\mathbf{x}}\log[p_1(\mathbf{z}_i(t)-\mu_t(\mathbf{x}))])(\nabla_{\mathbf{x}}\log[p_1(\mathbf{z}_i(t)-\mu_t(\mathbf{x}))])^T\right\} \tag{15}$$

is Fisher's information matrix for the *measurement-certain* case. Assuming, therefore, that P_d does not vary with t (this is a reasonable approximation in the far-field tracking situation), it therefore follows that

$$\mathbf{J} = q_2\mathbf{J}^0 = q_2\sum_{t=1}^{T}\mathbf{J}_t^0 \tag{16}$$

with an appropriate *information-reduction factor* (less than unity) in the proportionality to account for the estimation algorithm's need to weigh which of its observations (if any) are relevant and which are spurious. Calculation of this proportionality constant is remarkably involved, unfortunately, arising from the need to evaluate a high-dimensional integral. However, the proportionality constant has been tabulated for a number of cases, and since it is sparsely parametrized this is sufficient for many needs. An example is shown in Table I, and there is some illustration in Figure 3. This would seem to be a remarkable result: the presence of spurious data in one's observation set affects observation efficiency in a *scalar* way.

5) Simulation Results: In narrowband sonar signal processing, different bands in the frequency domain are defined by an appropriate cell resolution and a center frequency about which these bands are located. The received signal is sampled and filtered in these bands before applying an FFT and beamforming. The signal processor was assumed to consist of the frequency band $[500\text{Hz},\ 1000\text{Hz}]$ with a 2048-point FFT, so the frequency resolution (cell) is

$$C_\gamma = 500/2048 = 0.25 \text{ Hz} \tag{17}$$

For the bearing measurements, we assume that the sonar has 60 equal beams, resulting in an azimuth cell C_θ

$$C_\theta = 180^\circ/60 = 3.0^\circ \tag{18}$$

Assuming a uniform distribution within a cell the frequency and bearing measurements have standard deviations:

$$\begin{aligned} \sigma_\gamma &= 0.25/\sqrt{12} = 0.07 \text{ Hz} \\ \sigma_\theta &= 3.0/\sqrt{12} = 0.87^\circ \end{aligned} \tag{19}$$

We take the surveillance regions for bearing and frequency as

$$\begin{aligned} V_\theta &= [-20^\circ,\ 20^\circ] \\ V_\gamma &= [747\text{Hz},\ 753\text{Hz}] \end{aligned} \tag{20}$$

We restrict the validation gate to $g = 5$. In our simulations we take as ground truth that the target moves at $10m/s$ heading west and $5m/s$ heading north, starting from $(5000m, 35000m)$. The emitted frequency is 750Hz, so the true target parameter is $x = [5000\ 35000\ \ -10\ 5\ 750]$. There are 30 measurements, one each 30 seconds, for a total observation interval of 15 minutes. For the first 15 scans the platform moves in the northwest direction, and for the next 15, northeast, both at the speed of $7.1\ m/s$.

We examine the case where the noise corrupting the target-generated measurement is Johnson. In figure 1 are shown for a particular Johnson parameter value the configuration and results of estimation, along with the theoretically-predicted covariances. These latter are presented as two ellipses referring to the 99% confidence regions of the position estimates at the initial and final sampling instants. We can see that in 98 out of 100 Monte Carlo runs, the estimated initial and final positions fall into the 99% confidence region.

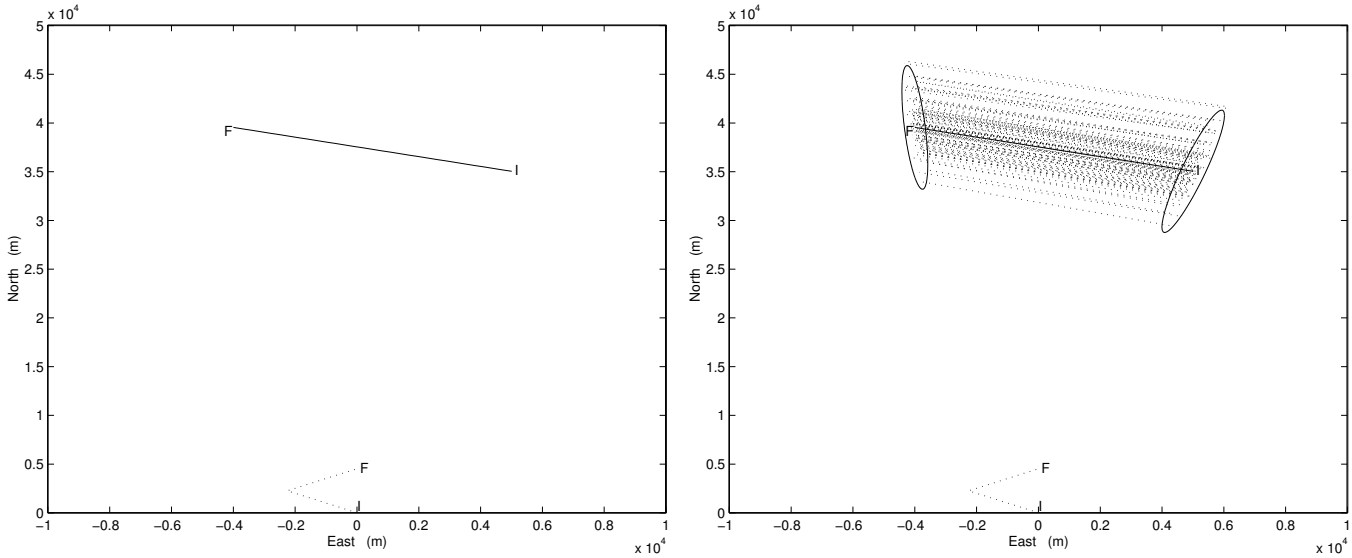

Fig. 1. Left: In corroborative simulation, the trajectories of target (solid) and platform (dotted). "I": Initial position of trajectories. "F": final position of trajectories. Right: The true and estimated trajectories from 100 Monte Carlo runs for Johnson noise case,$P_d = 0.9$,$\lambda V_g = 0.1$ and kurtosis = 11.5. Note that in almost all cases the estimated values lie within their respective ellipses.

B. The Dynamic Case: A CRLB for Target Tracking in Clutter

1) Introduction: Recently several papers have given clever recursively computable lower bounds for both linear and nonlinear stochastic dynamic systems, i.e., systems where the state is modeled as a random process evolving according to a stochastic equation. The approach by Tichavský *et al.* [35] presents the CRLB for the discrete-time multidimensional nonlinear state filtering problem in the form of a Riccati-like equation; while this is, in our opinion, a *major* contribution, it does not address the case with measurement origin uncertainty.

Here the CRLB is derived for the state of a linear dynamical system driven by and observed in the presence of additive white Gaussian noise (AWGN) and also with measurement origin uncertainty. Consequently, what is reported here can be considered a natural marriage between the results in [17], [18], [24] and those in [35][1]. It will be shown that the CRLB obeys a Riccati-like equation similar to that in [35], with the exception that the measurement-noise covariance term is multiplied by an IRF. The calculation of the IRF and the existence of efficient estimators are also addressed. These results are applicable to practically all the real-world target tracking problems with SNR under 20dB, target detection probability $P_d < 1$ and false alarm probability $P_{fa} > 0$, i.e., the cases in which there will be measurement origin uncertainty. Further, the CRLB is compared via simulation to the estimation MSE of two common target tracking algorithms, the probabilistic data association filter (PDAF) and the multi-frame (N-D) assignment algorithm. Details of this material are in [37].

2) The CRLB for Dynamic Systems without Measurement Origin Uncertainty: Consider the estimation of the state of a nonlinear dynamical system

$$x_{n+1} = f_n(x_n, v_n) \tag{21}$$
$$z_n = h_n(x_n, w_n) \tag{22}$$

where x_n is the system state at time n, $\{z_n\}$ is the measurement process, $\{w_n\}$ and $\{v_n\}$ are independent white processes, and f_n and h_n are (in general) nonlinear functions. Note that this problem has noisy dynamics (21), noisy measurements (22), but no measurement origin uncertainty — the measurement z_n is known to correspond to the state x_n.

The CRLB for this problem is given by the Fisher Information matrix sequence $\{J_n\}$ with the following proposition [35].

Proposition 1 *The sequence $\{J_n\}$ of Fisher information sub-matrices for estimating state vectors $\{x_n\}$ obeys the recursion:*

$$J_{n+1} = D_n^{22} - D_n^{21}(J_n + D_n^{11})^{-1} D_n^{12} \tag{23}$$

where

$$D_n^{11} = E\{-\nabla_{x_n}\nabla_{x_n}^T \ln\ p(x_{n+1}|x_n)\} \tag{24}$$
$$D_n^{12} = E\{-\nabla_{x_{n+1}}\nabla_{x_n}^T \ln\ p(x_{n+1}|x_n)\} \tag{25}$$
$$D_n^{21} = E\{-\nabla_{x_n}\nabla_{x_{n+1}}^T \ln\ p(x_{n+1}|x_n)\} = [D_n^{12}]^T \tag{26}$$
$$D_n^{22} = E\{-\nabla_{x_{n+1}}\nabla_{x_{n+1}}^T \ln\ p(x_{n+1}|x_n)\} + E\{-\nabla_{x_{n+1}}\nabla_{x_{n+1}}^T \ln\ p(z_{n+1}|x_{n+1})\} \tag{27}$$

The above CRLB has the same form as the Riccati equation of the linear dynamic system [1]. The Riccati-like form of the above system is clear.

3) Derivation of CRLB for Single Target Tracking: The model for single target tracking is similar to that of estimation for a linear dynamical system; that is, we have the model

$$x(k+1) = F(k)x(k) + v(k) \tag{28}$$
$$z(k) = H(k)x(k) + w(k) \tag{29}$$

where $x(k)$ is the system state at time k, $z(k)$ is the *true* measurement at time k, $v(k)$ and $w(k)$ are independent white Gaussian processes with zero mean and invertible covariance matrices Q and R. The difference is that there is uncertainty as to measurement origin, and implicit to this there is a variable number of measurements available at each scan. Now, from (23) – (27), it can be seen that only D_n^{22} is affected by the measurement uncertainty. Consequently, the CRLB for filtering with measurement uncertainty is available if we substitute the pdf of the measurement (conditioned on the state vector) in (27) by the corresponding pdf for the case with measurement uncertainty. Note that the expectation is now over $Z(k)$ (the observation set containing at most one true measurement), $x(k)$ and m_k (the number of observations).

Define

$$\begin{aligned} J_Z &\triangleq E\{-\nabla_{x(k)}\nabla_{x(k)}^T \ln\ p(Z(k)|x(k))\} \\ &= E\{(\nabla_{x(k)} \ln\ p(Z(k)|x(k)))(\nabla_{x(k)} \ln\ p(Z(k)|x(k)))^T\} \end{aligned} \tag{30}$$

[1]Our results do not explicitly deal with the case where, in addition to the measurement uncertainty, the dynamic system is nonlinear or the noises are non-Gaussian. Judging from [35] this would be a straightforward but notationally messy extension.

The expectation above is over $Z(k)$, $x(k)$ and m_k. Since the number of observations m_k *is known* to the estimator, we can write the expectation as

$$\begin{aligned} J_Z &= \sum_{m_k=1}^{\infty} P(m_k) E\{(\nabla_{x(k)} \ln\ p(Z(k)|x(k), m_k))(\nabla_{x(k)} \ln\ p(Z(k)|x(k), m_k))^T | m_k\} \\ &= \sum_{m_k=1}^{\infty} P(m_k) E\{(\nabla_{x(k)} \ln\ p(\{z_i(k)\}_{i=1}^{m_k}|x(k)))(\nabla_{x(k)} \ln\ p(\{z_i(k)\}_{i=1}^{m_k}|x(k)))^T\} \\ &= \sum_{m_k=1}^{\infty} P(m_k) J_Z(m_k) \end{aligned} \tag{31}$$

where

$$J_Z(m_k) \triangleq E\{(\nabla_{x(k)} \ln\ p(\{z_i(k)\}_{i=1}^{m_k}|x(k)))(\nabla_{x(k)} \ln\ p(\{z_i(k)\}_{i=1}^{m_k}|x(k)))^T\} \tag{32}$$

and

$$p(\{z_i(k)\}_{i=1}^{m_k}|x(k)) = \left[\prod_{j=1}^{m_k} p_0(z_j(k))\right] \left[(1-\epsilon(m_k)) + \frac{\epsilon(m_k)}{m_k} \sum_{i=1}^{m_k} \frac{p_1(z_i(k))}{p_0(z_i(k))}\right] \tag{33}$$

where $p_0(\cdot)$ and $p_1(\cdot)$ are the spatial pdf's of false (uniform) and true (Gaussian with covariance R) measurements. Note that the expectation in (32) is the expectation over $Z(k)$ and $x(k)$, conditioned on m_k[2]. We thus have

$$\begin{aligned} &p(\{z_i(k)\}_{i=1}^{m_k}|x(k)) \ = \\ &\frac{1}{V^{m_k}} \left[(1-\epsilon(m_k)) + \frac{\epsilon(m_k)V}{m_k} \frac{1}{\sqrt{|2\pi R|}} \sum_{i=1}^{m_k} e^{-\frac{1}{2}(z_i(k)-Hx(k))^T R^{-1}(z_i(k)-Hx(k))}\right] \end{aligned} \tag{34}$$

Now, define

$$\begin{aligned} J_Z^0 &= E\{(\nabla_{x(k)} \ln\ [p_1(z(k))])(\nabla_{x(k)} \ln\ [p_1(z(k))])^T\} \\ &= E\{[H^T R^{-1}(z(k) - Hx(k))][H^T R^{-1}(z(k) - Hx(k))]^T\} \\ &= E\{H^T R^{-1}(z(k) - Hx(k))(z(k) - Hx(k))^T R^{-1} H\} \\ &= H^T R^{-1} H \end{aligned} \tag{35}$$

as the counterpart of J_Z when there is no measurement origin uncertainty. The FIM in clutter, J_Z, is related to J_Z^0, the FIM in a "clean (or clear) environment," by the following proposition.

Proposition 2 *The FIM in clutter is given by*

$$J_Z = q_2(P_d, \lambda, R, V) J_Z^0 = q_2(P_d, \lambda, R, V) H^T R^{-1} H \tag{36}$$

where $q_2(P_d, \lambda, R, V)$ is a constant scalar dependent on the probability of detection, the false alarm density, the covariance of observation noise and on the volume of the observation region.

Ultimately, by combining Proposition 1, Proposition 2 and the assumed system model, the state estimation Fisher information matrix has the following convenient recursive form

$$J_{n+1} = Q^{-1} + q_2(P_d, \lambda, R, V) H^T R^{-1} H - Q^{-1} F (J_n + F^T Q^{-1} F)^{-1} F^T Q^{-1} \tag{37}$$

In the case that $q_2 = 1$ (i.e. no measurement origin uncertainty) this is recognizable as an expression for the evolution of the *inverse* of the Kalman filter estimation error, $P_{k|k}^{-1}$. This is, not unnaturally, in the form of the "information filter" [2].

Apart from the simplicity of the result, one thing that is interesting about equation (37) is that measurement origin uncertainty does not cause instability in the estimation. Nevertheless, in the measurement origin uncertain case all practical algorithms eventually lose track with probability one [19]. This is so since, with probability one, there will eventually be a sequence of missed detections and likely-appearing false alarms that the estimator cannot recover from. However, it must be recalled that all algorithms are suboptimal: the CRLB is the (not necessarily attainable) bound on the performance of an optimal estimator that operates on a batch of data.

[2]The sum over m_k begins with $m_k = 1$ because $m_k = 0$ means no observations at scan k, which gives no information.

4) The existence of efficient estimates: With the Cramér-Rao Lower Bound, an *efficient estimate* is defined as an estimate which satisfies the bound with equality. It is proven in [36] that, for the estimation of a parameter modelled as a random variable (the state of a noisy dynamic system falls in this category), a sufficient and necessary condition for efficiency is that the a posteriori probability density of the state must be Gaussian for all observations: that is, $p(x(k)|Z(k))$ must be Gaussian.

In target tracking with measurement origin uncertainty,

$$\begin{aligned} p(x(k)|Z(k), m_k) &= \frac{1}{c} p(Z(k)|x(k), m_k) p(x(k)|m_k) \\ &= \frac{1}{c} p(Z(k)|x(k), m_k) p(x(k)) \end{aligned} \tag{38}$$

where

$$\begin{aligned} p(Z(k)|x(k), m_k) &= p(Z(k)|m_k, \text{all false}, x(k)) p(\text{all false}|m_k) + \\ &\quad p(Z(k)|m_k, \text{one true detection}, x(k)) p(\text{one true detection}|m_k) \\ &= \left[\prod_{j=1}^{m_k} p_0(z_j(k)) \right] \left[(1 - \epsilon(m_k)) + \frac{\epsilon(m_k)}{m_k} \sum_{i=1}^{m_k} \frac{p_1(z_i(k))}{p_0(z_i(k))} \right] \end{aligned} \tag{39}$$

In target tracking, the locations of false alarms are modeled as having a uniform distribution over the observation volume [2]. Obviously, in the above equation, $p(Z(k)|x(k), m_k)$ is not Gaussian (it is a uniform-Gaussian mixture) except in trivial cases in which there are neither missed detections nor false alarms: this means $p(x(k)|Z(k), m_k)$ is not Gaussian. It is perhaps surprising that in the general target tracking scenario, efficient estimators do not exist, but the above discussion clearly indicates why.

When there are no false alarms, the target tracking problem has measurement-origin uncertainty from the non-unity probability of detection only. In this case, the only possible values of m_k are 0 (the target is not detected, thus no measurements) and 1 (the target is detected and the measurement is the true measurement). It can be easily shown that both $p(x(k)|Z(k), m_k = 0)$ and $p(x(k)|Z(k), m_k = 1)$ are Gaussian in the linear Gaussian model as (28) and (29). The bound for fixed m_k is thus approachable, i.e., the bound given by $J_Z(m_k)^{-1}$ can be reached. However, what we are interested in is the bound for the *average performance*, which is a weighted average of $J_Z(m_k)^{-1}$ over all possible m_k's, as in (37). Since we have the discrete random variable m_k, $p(x(k)|Z(k), m_k)$ is no longer Gaussian, and unfortunately, there is no efficient estimator even for the case of no false alarms. This is illustrated in an example.

5) Numerical Results: We consider a kinematic target moving in a two-dimensional space, as in (28) and (29). The state vector in Cartesian coordinates is

$$x = [\xi \; \dot{\xi} \; \eta \; \dot{\eta}] \tag{40}$$

The state and observation noise covariance matrices Q and R are given by

$$Q = \begin{bmatrix} \frac{1}{3}T^3 & \frac{1}{2}T^2 & 0 & 0 \\ \frac{1}{2}T^2 & T & 0 & 0 \\ 0 & 0 & \frac{1}{3}T^3 & \frac{1}{2}T^2 \\ 0 & 0 & \frac{1}{2}T^2 & T \end{bmatrix} \tilde{q} \qquad R = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \sigma_w^2 \tag{41}$$

where T is the sampling time, $\tilde{q}$ is the state noise power spectral density and σ_w^2 is the measurement noise variance. The F and H matrices are given by

$$F = \begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{42}$$

The observation is of position only, and is thus a two-dimensional vector; more detailed discussion of single target tracking models can be found in [1], [2]. In the figures that follow, the CRLBs shown are for position, meaning that the $(1, 1)$ element of J^{-1} is shown and plotted. Also, it should be noted that *only the CRLB steady-state values* are shown. Finally, we note that in the computation of the (scalar) IRF, q_2, we use a finite observation volume, in this case $V = 400$ in two dimensions. A finite V is not necessary, but one that is sufficiently large does not detract, and can be helpful in the IRF's numerical computation.

C. The CRLB vs. System Parameters

Table I shows q_2 values for the case $\tilde{q} = 1$ and $\sigma_w^2 = 1$. As can be seen, even in quite heavy clutter and for high missed-detection rates, the information lost appears not too large. One aspect that is interesting and intuitive is that the loss of a detection becomes more unfortunate as the clutter becomes more dense. For example, consider the $\lambda = 0.008$ row from Table I: one might expect that the decrease in probability of detection from 80% to 70% would scale q_2 by a factor of $7/8 = 87.5\%$

	$P_D = 0.7$	$P_D = 0.8$	$P_D = 0.9$
$\lambda = 0.0001$	0.6814	0.7700	0.8890
$\lambda = 0.0020$	0.6341	0.7472	0.8491
$\lambda = 0.0040$	0.5722	0.6860	0.8144
$\lambda = 0.0060$	0.5460	0.6498	0.7727
$\lambda = 0.0080$	0.5257	0.6290	0.7457
$\lambda = 0.0100$	0.5037	0.6096	0.7367
$\lambda = 0.0120$	0.4679	0.5729	0.6938
$\lambda = 0.0140$	0.4479	0.5343	0.6801
$\lambda = 0.0160$	0.4356	0.5330	0.6590
$\lambda = 0.0180$	0.4180	0.5133	0.6312

TABLE I

VALUE OF q_2 AT DIFFERENT FALSE ALARM DENSITIES AND PROBABILITIES OF DETECTION. THE STATE AND MEASUREMENT NOISES ARE SET AS $\tilde{q} = 1$ AND $\sigma_w^2 = 1$.

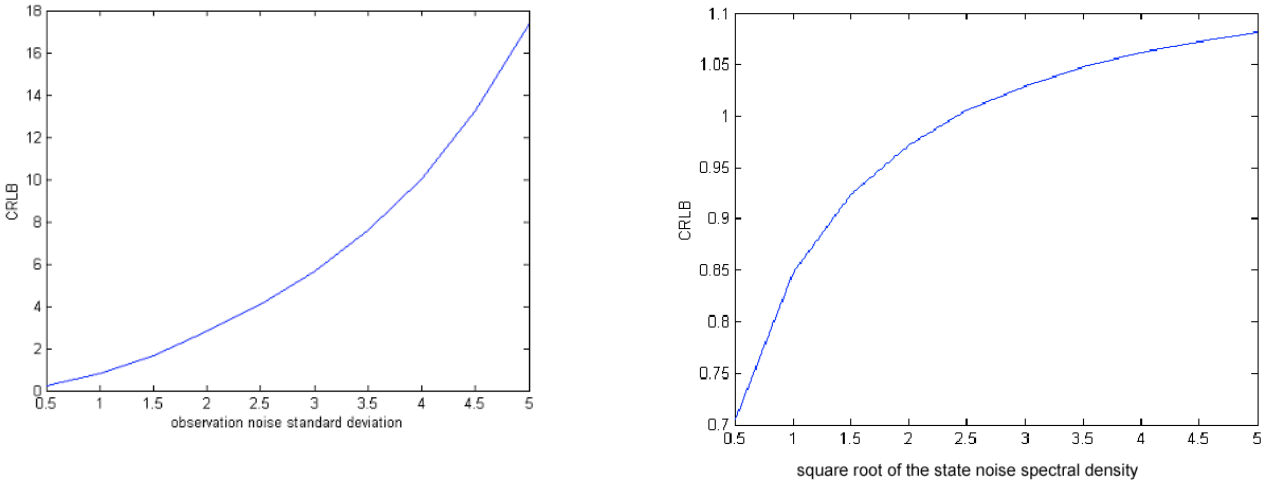

Fig. 2. Left: CRLB vs measurement noise standard deviation σ_w, at state noise spectral density $\tilde{q} = 1$, probability of detection $P_d = 90\%$ and clutter density $\lambda = 10^{-4}$. Right: CRLB vs $\sqrt{\tilde{q}}$, the square root of the state noise spectral density, at measurement noise variance $\sigma_w^2 = 1$, probability of detection $P_d = 90\%$ and clutter density $\lambda = 2 \times 10^{-3}$.

since only this portion of the original valid measurements remain, but in fact the factor is $0.5257/0.6290 = 83.6\%$. Further, at double the clutter, the factor is further reduced, to 81.7% — clearly the "confusion" caused by the clutter exacerbates the loss of valid measurements.

The left plot of Figure 2 shows the effect of the measurement noise standard deviation on the CRLB: there appears to be a quadratic relationship, as in the case without measurement origin uncertainty. The right plot similarly suggests the relationship between the CRLB and $\sqrt{\tilde{q}}$, the square root of the state noise power spectral density. For $\sqrt{\tilde{q}} = 0$ the CRLB is zero, since this corresponds to "straight-line" motion and filtering amounts to estimation of initial parameters of position and velocity, which in steady state will have perfect estimates. As expected, the CRLB rises sharply from zero, but increases more slowly as the trajectory becomes more random (unpredictable). However, note that without measurement origin uncertainty this CRLB would be limited by σ_w^2, since each measurement would yield position information of this accuracy; *with* measurement origin uncertainty the CRLB continues to increase, and apparently does not reach an asymptotic value.

Figure 3 deals explicitly with measurement origin uncertainty, and show, respectively, the effects of the clutter density λ and of the probability of detection P_d on the CRLB. What is interesting here is that the sensitivity to clutter (as long as there is *some* clutter) is not great, but that the estimation degradation caused by missed detections is highly significant.

D. Performance of the PDAF and N-D Assignment vs. CRLB

We compare the performances of the classic Probabilistic Data Association Filter (PDAF) [2], and of the 2-D, 3-D and 4-D (multi-scan) assignment algorithms [29], to the CRLB; only errors from tracks that are not lost are reported. Here the dimensionality of the target has been reduced from two to one, so that lost track effects are more easily controlled. The assignment algorithms operate on sliding windows of, respectively, one, two and three scans of data[3]. They evaluate, via a

[3]The 2-D, 3-D and 4-D terminology arises from the assignment of measurements to tracks. That is, a 3-D assignment algorithm associates tracks to two scans of observations, for a total of three "lists."

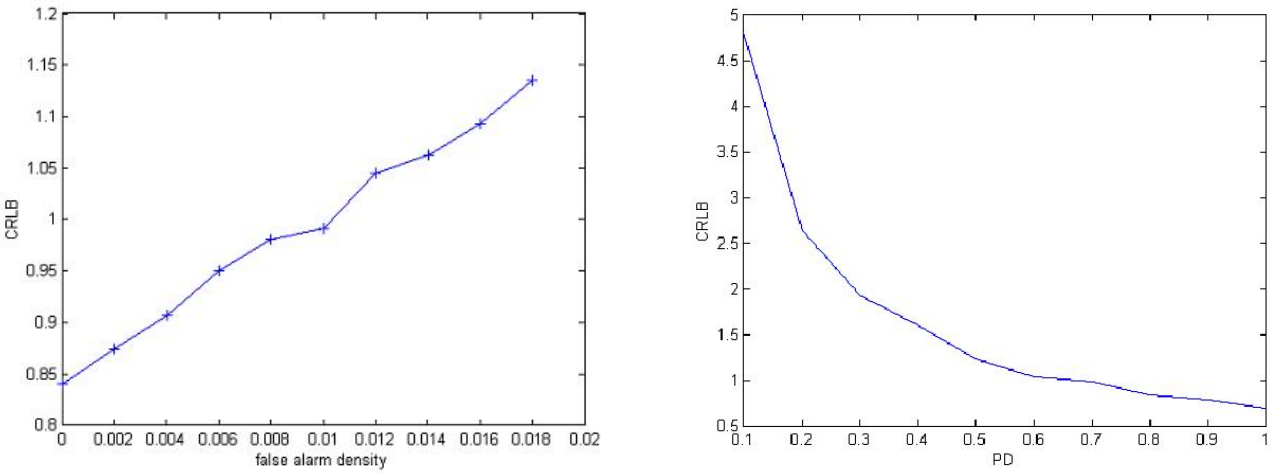

Fig. 3. Left: CRLB vs clutter density λ, at state noise spectral density $\tilde{q} = 1$, measurement noise variance $\sigma_w^2 = 1$ and probability of detection $P_d = 90\%$. Right: CRLB vs probability of detection P_d, at state noise spectral density $\tilde{q} = 1$, measurement noise variance $\sigma_w^2 = 1$ and clutter density $\lambda = 10^{-5}$.

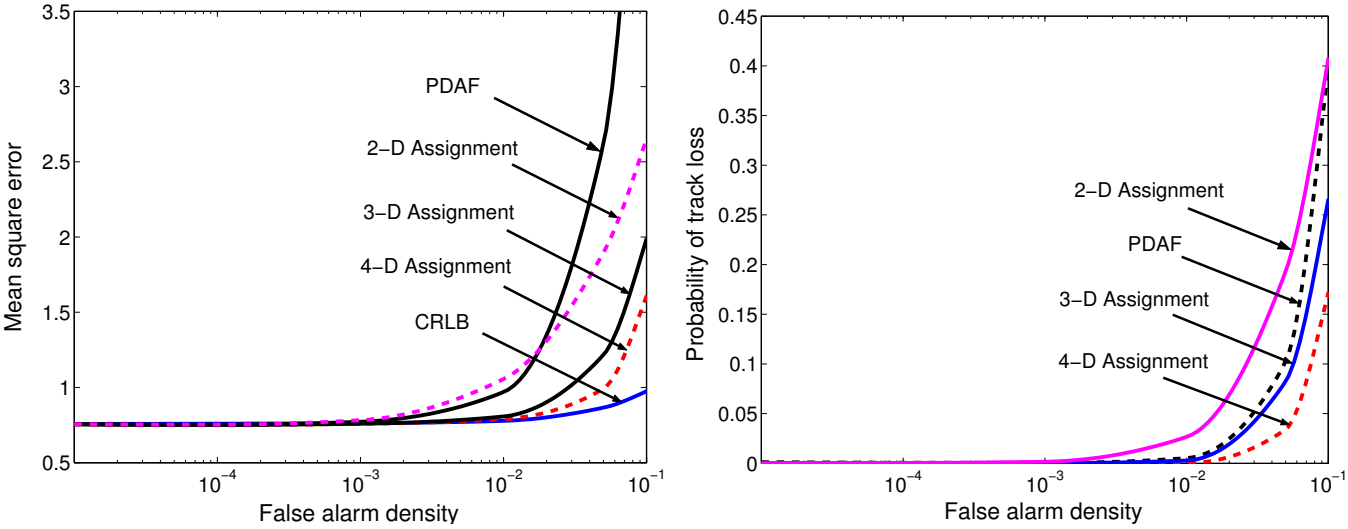

Fig. 4. Left: CRLB compared to MSE of 2-D, 3-D and 4-D assignment and of PDAF, with $P_d = 1$ at different false alarm densities. Only MSE's from those tracks that are not lost are considered. Right: Probability of track loss, for 2-D, 3-D and 4-D assignment and PDAF, with $P_d = 1$ at different false alarm densities.

Kalman filter, the likelihood of each possible association combination being true, and choose the most likely one. The PDAF generates an optimal state estimate at each scan of data under the assumption that all previous data has given rise to a Gaussian posterior — in fact, the "forcing" of such a moment-matched posterior probability density function is a key feature of the PDAF. Note: both of these algorithms are sub-optimal.[4]

In all the following simulations, we set the state noise spectral density $\tilde{q} = 1$ and the observation noise variance $\sigma_w^2 = 1$. The performances of the PDAF and assignment algorithms relative to the CRLB are shown on the left in Figure 4, for the case that the probability of detection $P_d = 100\%$. All trackers appear to follow the CRLB up to a threshold level of the false alarm density. For the 2-D assignment algorithm and the PDAF, this threshold is somewhat lower than for the 3-D and 4-D trackers. It is interesting that at low clutter densities (i.e. low λ) the PDAF has lower MSE than 2-D assignment algorithm; while at higher λ, 2-D assignment algorithm is apparently better. However, it must be recalled that lost tracks[5] are ignored here, and hence the PDAF's good tracking ability is to some extent penalized: in fact the PDAF loses fewer tracks, as shown in the right plot. The two multi-scan trackers (3-D and 4-D) each have significantly higher clutter thresholds, and in fact remain

[4]While assignment chooses the most likely association, this is not necessarily the true one; however, the state covariance matrices are propagated as if this association was the correct one.

[5]A track is deemed "lost" when the error in the position estimate is larger than four times the measurement noise standard deviation for three scans in a row (after this there is practically no chance of "recovery").

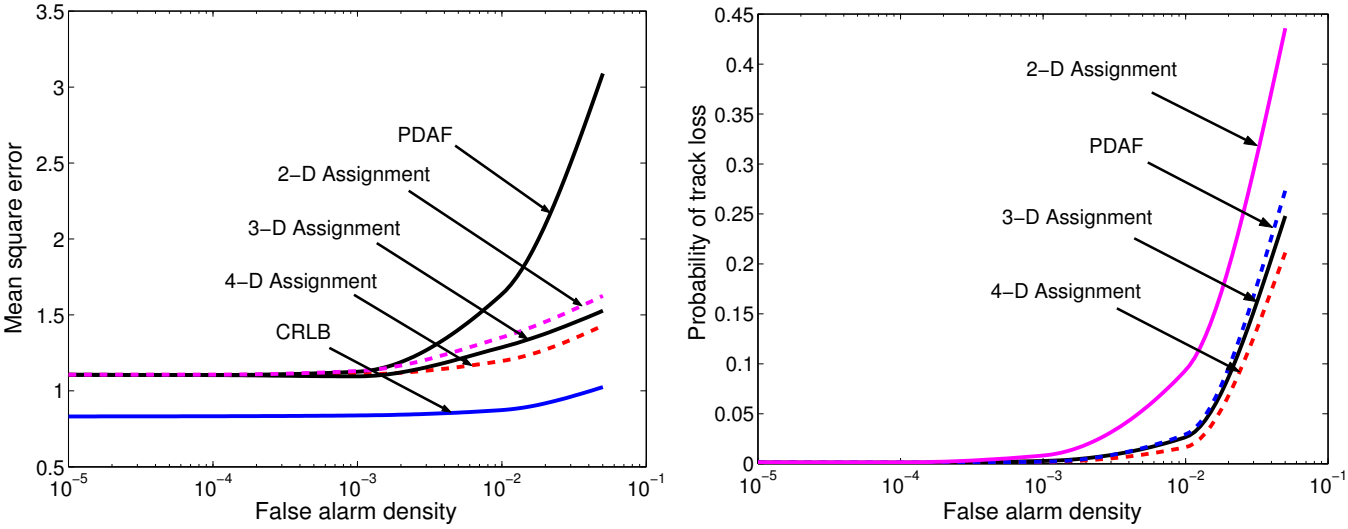

Fig. 5. Left: CRLB compared to MSE of 2-D, 3-D and 4-D assignment and of PDAF, with $P_d = 0.9$ at different false alarm densities. Only MSE's from those tracks that are not lost are considered. Right: Probability of track loss, for 2-D, 3-D and 4-D assignment and PDAF, with $P_d = 0.9$ at different false alarm densities.

quite close to the CRLB until the clutter becomes heavy.

Figure 4 is repeated in Figure 5, but now with a 90% probability of detection. The differences between these figures is interesting: even for low clutter, all trackers have significantly higher error than the CRLB (which is not necessarily achievable). From this figure we additionally see a marked increase in the degree of MSE inflation of the PDAF; and, interestingly, the multi-scan assignment algorithms considered have comparable performance. Here we see that the PDAF is actually comparable to the first multi-scan algorithm (3-D) in terms of lost tracks: it loses substantially fewer tracks than does 2-D assignment algorithm (i.e. a "nearest-neighbor" filter). This explains the disappointing MSE performance of the PDAF in Figure 5: the PDAF keeps more tracks than does 2-D assignment algorithm, but in doing so its accuracy suffers.[6]

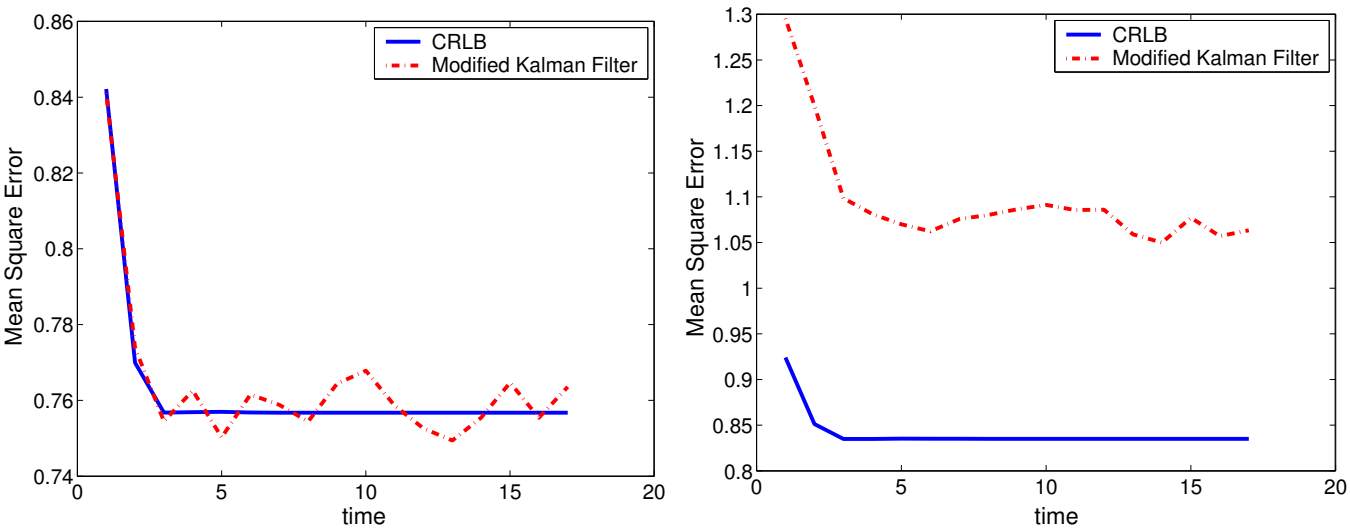

Fig. 6. Left: Kalman Filter MSE versus CRLB with $P_d = 1$ and no false alarms. Right, same for $P_d = 0.9$.

1) Examples of Efficiency and Lack of Efficiency of Estimators: Next we examine the case that there are no false alarms, and in which the only uncertainty comes from the non-unity probability of detection. The CRLB and the mean square error given by a Kalman Filter with $P_d = 1$ are shown on the left in Figure 6; here the filters and CRLB evaluation are initialized according to the two-point differencing scheme of [1]. It can be seen that the Kalman Filter approaches the CRLB, and that it is an efficient estimator. The right plot shows the comparison with $P_d = 0.9$. Since traditional Kalman Filter does not work

[6]Two possible reasons for this are that in keeping more tracks, it keeps tracks that are intrinsically "tough" and hence have higher MSE; and also that the PDAF's good in-track performance relies on a continual adaptation of its "bandwidth" to match current conditions, which amounts to detuning.

with non-unity probability of detection, we use a "modified" Kalman Filter. When there are observations, the modified Kalman filter works just like traditional Kalman Filter. When there are missed detections, i.e., there is no observations, the modified Kalman Filter uses the predicted measurement from the previous measurements to propagate the states. This modified Kalman Filter is optimal here, and there is no concern about lost tracks. However, as can be seen, even it cannot approach the CRLB, and hence *there is no efficient estimator* even in this simplest of cases that involves measurement origin uncertainty (in this case, the posterior density is a Gaussian mixture). That is, for an efficient (CRLB-reaching) estimator to exist it must be that the ensemble-averaged posterior error has a Gaussian probability density; and when there is measurement origin uncertainty it is impossible that this can be so.

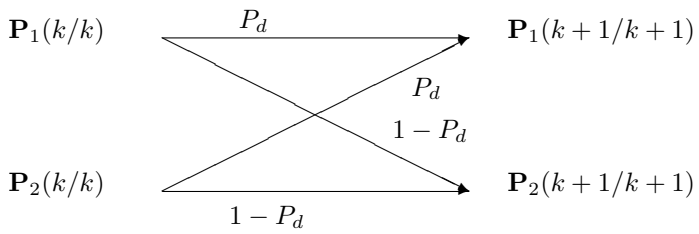

Fig. 7. An illustration of the HYCA procedure for finding the steady-state covariance for a Kalman filter operating in a stochastic environment with missed detections only.

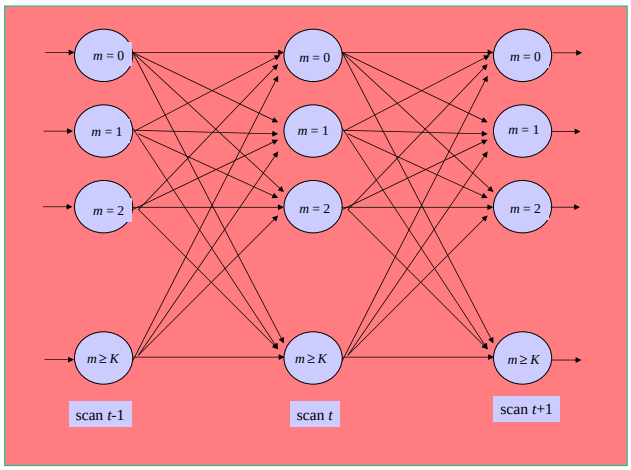

Fig. 8. More general HYCA procedure. In this case m refers to the number of gated measurements at each scan. When this number becomes sufficiently large (say, $m_k = 10$) the track may be considered lost, since the probability of continued increase in m_{k+1} is high.

2) Hybrid Conditional Averaging: The previous CRLB technique is exact and analytic. Its weaknesses are the looseness of it bound and the fact that it says nothing about the in-track performance. An analytic technique that accomplishes this is hybrid conditional averaging (HYCA), a Markov chain analysis. We present that now, and for simplicity begin with the case of missed detections only.

The state estimation-error covariance matrix is updated by the Kalman filter as follows ([1]):

$$\begin{aligned}
&\mathbf{P}(k+1|k) = \mathbf{F}\mathbf{P}(k|k)\mathbf{F}' + \mathbf{Q} \\
&\mathbf{S}(k+1) = \mathbf{H}\mathbf{P}(k+1|k)\mathbf{H}' + \mathbf{R}(k) \\
&\mathbf{W}(k+1) = \mathbf{P}(k+1|k)\mathbf{H}'\mathbf{S}(k+1)^{-1} \\
&\mathbf{P}(k+1|k+1) = \mathbf{P}(k+1|k) - \mathbf{W}(k+1)\mathbf{S}(k+1)\mathbf{W}'(k+1)
\end{aligned} \tag{43}$$

With $\mathbf{R}(k) = \mathbf{R}$ this evolves to a steady state which characterizes the performance of the filter, and this solution can be computed from the associated algebraic Riccati equation (ARE). However, in the case that detections can be missed, the measurement noise has a stochastic covariance ($\mathbf{R}(k)$), and there *is* no steady-state, but a stationary matrix-valued random process. Therefore, it is legitimate to seek the *expected value of the error-covariance*.

Let us assume that the measurement noise covariance matrix can be modeled as:

$$\mathbf{R}(k) = \begin{cases} \mathbf{R} & detection \\ \infty\mathbf{I} & miss \end{cases} \tag{44}$$

The method here employed is the HYCA (hybrid conditional averaging) method of [19]. (For justification of the approach, and comparison of HYCA results to simulation, the reader is directed to [2].) Under this scheme we embed a binary random process $\{b(k)\}$ into the Riccati equation such that

$$b(k) = \begin{cases} 1 & detection \\ 2 & miss \end{cases} \qquad P\{b(k) = i\} = \begin{cases} P_d & i = 1 \\ 1 - P_d & i = 2 \end{cases} \tag{45}$$

Then, with $\mathbf{P}_{ij}(k+1|k+1)$ defined as the estimation error covariance given $b(k) = i$ and $b(k+1) = j$ (obtainable from (43) with $\mathbf{R}(k)$ according to (44)), we get

$$\mathbf{P}_j(k+1|k+1) = \sum_{i=1}^{2} \mathbf{P}_{ij}(k|k) P\{b(k) = i\} \tag{46}$$

The state probabilities at time k, $\mu_i(k) = P\{b(k) = i\}, i = 1, 2$ are derived from the state probabilities at time $k-1$ and the Markov transition matrix:

$$\mu(k) \triangleq \begin{bmatrix} \mu_1(k) \\ \mu_2(k) \end{bmatrix} = \Pi' \mu(k-1) \qquad \text{where} \qquad \Pi = [\pi_{ij}] = \begin{bmatrix} P_d & 1 - P_d \\ P_d & 1 - P_d \end{bmatrix} \tag{47}$$

The matrices $\mathbf{P}_i(k|k)$ reflect the state error covariance given that $s(k) = i$. The mean state covariance matrix (i.e. what we want) is given by

$$\mathbf{P}(k|k) = \sum_{j=1}^{2} \mathbf{P}_j(k|k) \mu_j(k) \tag{48}$$

This can be evaluated when this matrix Markov process reaches steady state.

Figure 7 shows a flow-diagram that reflects the mechanics of the iterations when the only measurement origin uncertainty is from missed detections. The more general procedure is illustrated in Figure 8; in this case an appropriate information reduction factor controls the iterations, since the expectation is over the measurement origin uncertainty that remains even when the number of gated detections is known. It is generally observed that an increased number of gated measurements result in a larger posterior covariance. This results in a larger gate, which in turn increases the number of measurements; that is, a "runaway" occurs in which all tracks are lost with probability one, as is in practice true.

II. Track Testing for Fluctuating Targets

A. Background

In active sonar tracking applications, targets can fade: the target's detection probability can shift suddenly between high and low values. We examine the performance of track management (confirmation and termination) routines where target detections are based on an underlying Hidden Markov Model (HMM) with high and low detection states. Rule-based track confirmation tests are compared including M/N rules and rules that differentiate the measurements as to receiver source (M/N from at least C sensors), each of which is sub-optimal compared to a fixed length likelihood ratio test. We show that significant performance improvements (to near-optimal) can be obtained using a composite track confirmation test that combines two or three such rules in a logical OR operation. Track termination tests are next compared and it is shown that a Bayesian sequential test (the Shiryaev test) yields dramatic performance improvements over a K/N track termination rule and the Page test. The model-based results are validated using simulations of a multistatic tracking scenario.

We consider a track management model that has characteristics of both centralized and decentralized tracking systems. Specifically, a centralized track management model is used that processes time ordered measurements from all sensors and include sensor origin information. Recognizing that sensor detection probability can be in a high or low state, the centralized model will apply track management logic that accounts for this variable sensor performance by considering measurement sensor origin. The sensor detection performance is modeled as a two-state Markov chain with high and low detection states. Target-originated measurements (binary detection events) can therefore be described using a Hidden Markov Model (HMM) structure. While there is a well-developed base of literature covering track confirmation and termination for sensors with a fixed probability of detection (P_d) of the target on a single scan, research pertaining to the track management problem for sensors with P_d based on a Markov model has only recently been considered [8].

Some of the major track confirmation routines using a fixed P_d detection model include [2], [4]:

1) M/N tests (a track is confirmed if at least M detections are received over N scans of data). Performance of M/N tests has typically been analyzed by modeling the detection sequence as a Markov chain and by accounting for a variable probability of false alarm (P_{fa}) based on the size of the track gate validation region.

2) Sequential Probability Ratio Tests (SPRT). These tests compare the probabilities of receiving the measurement sequence under the true target and false alarm hypotheses and can consider binary detection events, the kinematic data (measurement innovations) and the amplitude of detections. These tests can also be used as a track score function.
3) Bayesian sequential tests. These tests are similar to the SPRT tests, but which apply Bayes' rule to recursively update the posterior probabilities of each hypothesis.
4) Batch techniques. These techniques process detections over multiple frames in a batch algorithm such as in a Track-Before-Detect system. Another example is the use of the Hough transform technique [34].

Similar to track confirmation routines, track termination routines can take the following forms [2], [4]:

1) K/N tests (a track is terminated if K or fewer detections are received in the last N scans).
2) Track score tests. These may include the SPRT or Bayesian sequential tests. If the track score (related to the probability that the detection sequence is the result of a true track) falls below a certain value, the track is terminated.

Thus we discuss the performance of K/N-based and sequential track termination tests when target-originated measurements are described by a HMM with high and low P_d Markov states. Using only the binary detection events, it is shown that the K/N test outperforms the Page test over a portion of its operating characteristic region. This result is surprising considering the fact that the Page test is proven to be the optimal sequential test for quickest detection of a change in measurement distribution and we show how when the HMM-based detection statistics are used, a key assumption in the optimality proof for the Page test is no longer satisfied. It is next shown that by using a Bayesian version of a sequential test (the Shiryaev test), significant performance improvement is obtained compared to the K/N test.

B. Track Confirmation

Given the detection sequence δ_1^k, the optimum hypothesis test for track confirmation is given by a Likelihood Ratio Test (LRT) using the Neyman-Pearson lemma [36] (note that this test may no longer be optimum when additional information, such as kinematic or amplitude information is used). Consider the hypotheses H_1 (target present) and H_0 (target absent). We derive the LRT for the observed detection sequence, δ_1^k, given that a tentative track has been initiated (denoted by TT). All tentative tracks require that $\delta_1 = 1$. The likelihood ratio, Λ, therefore becomes

$$\Lambda = \frac{\Pr\{\delta_1^k | H_1, TT\}}{\Pr\{\delta_1^k | H_0, TT\}}. \tag{49}$$

1) Likelihood Function Under H_1: Use of conditional probability rules gives

$$\Pr\{\delta_1^k | H_1, TT\} = \frac{\Pr\{\delta_1^k, TT | H_1\}}{\Pr\{TT | H_1\}}. \tag{50}$$

where the denominator is a normalizing constant to ensure $\Pr\{\delta_1^k | H_1, TT\}$ is a proper probability mass function. For a given sensor the detection sequence is dependent upon its underlying Markov detection state. Given that detections are independent across sensors, application of the total probability theorem yields

$$\Pr\{\delta_1^k, TT | H_1\} = \prod_{i=1}^{N_s} \sum_{k_1=0}^{1} \cdots \sum_{k_n=0}^{1} \Pr\{\delta_{i1}, \ldots, \delta_{in} | H_1, u_{i1} = k_1, \ldots, u_{in} = k_n\} \Pr\{u_{i1} = k_1, \ldots, u_{in} = k_n\} \tag{51}$$

where u_{ij} and δ_{ij} are the Markov detection state and detection observation for the ith sensor on scan cycle j. Note that measurements are now tagged with sensor origin and $n = k/N_s$ represents the number of scan cycles in the data set. The underlying Markov detection process gives

$$\Pr\{u_{i1} = k_1, \ldots, u_{in} = k_n\} = \prod_{l=2}^{n} \Pr\{u_{il} = k_l | u_{i(l-1)} = k_{l-1}\} \Pr\{u_{i1} = k_1\}. \tag{52}$$

Finally by further considering detections to be independent across scan cycles for a given sensor (conditioned on the underlying Markov detection state) we obtain

$$\Pr\{\delta_1^k, TT | H_1\} = \prod_{i=1}^{N_s} \sum_{k_1=0}^{1} \cdots \sum_{k_n=0}^{1} \prod_{j=1}^{n} \Pr\{\delta_{ij} | H_1, u_{ij} = k_j\} \prod_{l=2}^{n} \Pr\{u_{il} = k_l | u_{i(l-1)} = k_{l-1}\} \Pr\{u_{i1} = k_1\}. \tag{53}$$

Further details are available in [5].

TABLE II
MULTISTATIC TRACKING SCENARIO PARAMETERS

Parameter	Value
Number of sensors, N_s	4
Number of scan cycles, n	3, 4, 5
Scan cycle period, T	60 s
Markov transition probabilities, p, q	0.1/3, 0.1
Detection probabilities, P_d^L, P_d^H	0.1, 0.9
False alarm density, λ	3.14×10^{-8} m^{-2}
Measurement noise covariance, $\mathbf{R}$	$\mathbf{I}_2\,(243.2\,\text{m})^2$
Prior velocity uncertainty, $\mathbf{P}_1^v$	$\mathbf{I}_2\,(1\,\text{m/s})^2$
Process noise parameter, $\tilde{q}$	$10^{-4}\,\text{m}^2/\text{s}^3$
Validation gate probability, P_G	99%

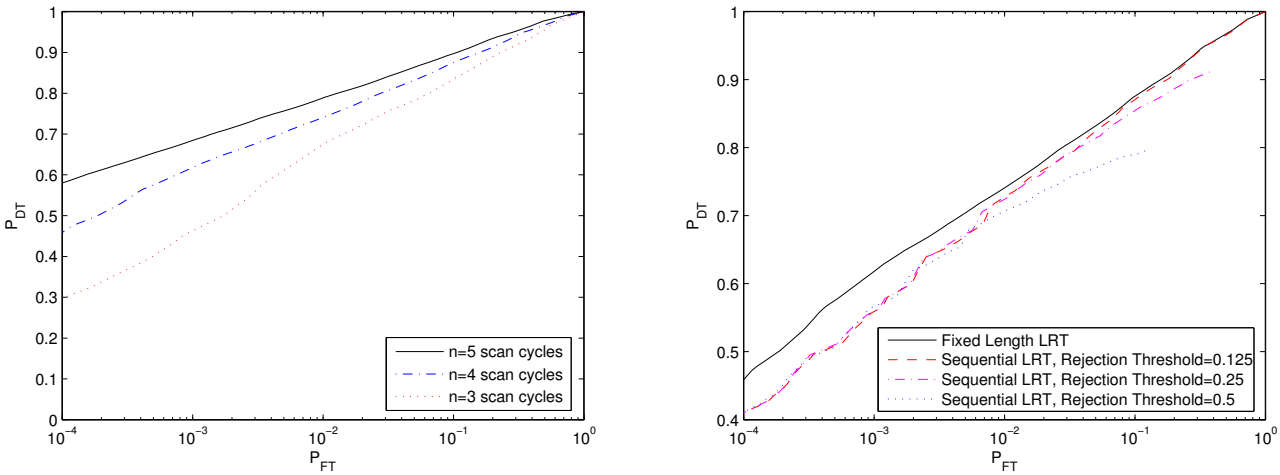

Fig. 9. Left: Track confirmation test performance using a fixed length LRT of n scan cycles. Right: Track confirmation test performance comparing a fixed length to a sequential LRT.

2) Likelihood Function Under H_0: Under H_0, detections are independent of the underlying Markov detection state and are only a function of $P_{fa}(\delta_1^k)$. Applying the conditional probability rules,

$$\Pr\{\delta_1^k|H_0, TT\} = \frac{\Pr\{\delta_1^k, TT|H_0\}}{\Pr\{TT|H_0\}} \tag{54}$$

where again the denominator is a normalizing constant to ensure $\Pr\{\delta_1^k|H_0, TT\}$ is a proper probability mass function.

False alarms are independent across sensors and scan cycles resulting in

$$\Pr\{\delta_1^k, TT|H_0\} = \prod_{i=1}^{N_s}\prod_{j=1}^{n}[P_{fa}]^{\delta_{ij}}[1-P_{fa}]^{1-\delta_{ij}} \tag{55}$$

where P_{fa} is a function of the detection sequence [5].

3) Likelihood Ratio Test for Track Confirmation: While use of a LRT for track confirmation may not be practical in tracking applications due to the computational complexity[7] in computing the LRT for a given detection sequence, this test provides an upper bound for the achievable performance of other track confirmation rulesets using the same information. As such this test is useful in evaluating other track confirmation methods. Two forms of the LRT test are available—fixed length and sequential.

In a fixed length LRT, a track confirmation decision is not made until detection data from all n scan cycles are obtained from all N_s sensors. If the likelihood ratio exceeds a threshold, α, the tentative track is confirmed. Otherwise it is rejected. The performance of the LRT can be viewed on a System Operating Characteristic (SOC), which plots the probability of accepting/detecting a true target track (P_{DT}) vs. the probability of accepting a false target track (P_{FT}), parameterized on α.

[7]The computational complexity is $\mathcal{O}(2^k)$ where k is the number of scans. For 4 sensors and 5 scan cycles, $k = 20$ and the complexity is $\mathcal{O}(10^6)$.

To compute the SOC, one evaluates the likelihood ratio for all possible detection sequences. To compute P_{DT} and P_{FT} at a given LRT threshold value,

$$P_{DT} = \sum_{\text{all sequences } \delta} \Pr\{\delta|H_1, TT\} I(\Lambda(\delta) > \alpha) \tag{56}$$

$$P_{FT} = \sum_{\text{all sequences } \delta} \Pr\{\delta|H_0, TT\} I(\Lambda(\delta) > \alpha) \tag{57}$$

where $I(\Lambda(\delta) > \alpha)$ is an indicator function such that probabilities are summed only for those sequences where a tentative track is accepted.

To evaluate the performance of the fixed length LRT as the number of scan cycles, n, is varied, a multistatic tracking problem is considered. Table II provides the parameters for this tracking problem. For the purposes of this study sensor measurements were considered to arrive sequentially at an interval of T/N_s. Fig. 9 (left) shows the SOC for the cases where $n = 3, 4, 5$ scan cycles. To construct this model-based performance plot, all combinations of detection sequences were enumerated and (P_{FT}, P_{DT}) points were computed according to (56) and (57) as a function of α. As expected, performance improves as more information (scan cycles) is used in the decision making process. Next, a sequential LRT test is considered. For this test, as each scan cycle is completed, an LRT test is performed with both a track acceptance and a track rejection threshold. Under this test, all tentative tracks that have not been accepted when the last scan cycle is completed are also rejected. While theory states that the performance of a sequential LRT with a given average duration can be no worse than that of a fixed length LRT with the same duration [32], one would expect that a sequential LRT with a maximum duration of n scan cycles would perform worse than a fixed length LRT of n scan cycles since the average duration of the sequential LRT is shorter than that of the fixed length LRT. Fig. 9 (right) compares fixed length LRT ($n = 4$) to sequential LRTs with different track rejection thresholds.

As the track rejection threshold is lowered, performance of the sequential LRT improves although in the limiting case it remains sub-optimal compared to the fixed length LRT. This is caused by a higher acceptance of false tracks when tracks are accepted in early scan cycles under the sequential LRT that would ultimately be rejected if information were allowed to accumulate for the full n scan cycles.

C. Track Termination

In considering the track termination module of a tracking system, we consider sequential tests. It is well known that the CUSUM (Page) test yields the quickest detection of a change of distribution for the case of i.i.d. observations [3]. In fact, in a (highly) simplified target tracking model where detection of target and false alarms can be described by Bernoulli random processes with fixed parameters (i.e., P_{fa} is independent of the detection sequence), the K/N track termination rule becomes a sufficient statistic for discrete Page test thresholds. The optimality results of the Page test have also been extended to some non-i.i.d distributions (including certain classes of Markov chain structures) [22]. However, there have been no global optimality results proven for the case where the distributions are characterized by HMMs except in the special cases considered by [15] whose conditions are not satisfied in this application.

In addition to examining the Page test, we also implement a stopping rule based on a Bayesian formulation of the problem called the Shiryaev rule [3] which has optimality properties similar to the Page test.

1) Page Test for HMMs: In an HMM the formulation for the Page test is given by [7]

$$s_k = \ln \frac{\Pr\{\delta_k|\delta_1^{k-1}, H_0\}}{\Pr\{\delta_k|\delta_1^{k-1}, H_1\}} = \ln \frac{f_0(\delta_k|\delta_1^{k-1})}{f_1(\delta_k|\delta_1^{k-1})} \tag{58}$$

$$c_k = \max(c_{k-1} + s_k, 0) \tag{59}$$

where with each scan c_k is compared to a threshold h and $c_0 = 0$. If the threshold is exceeded then the track is terminated. Note that the hypothesis convention used throughout this report where H_1 represents target present (pre-change distribution for track termination testing) and H_0 represents target absent (post-change distribution for track termination testing) is opposite of that used in the quickest detection literature where H_0 represents the pre-change distribution and H_1 the post-change distribution.

Under H_0 (target absent), the likelihood function becomes

$$f_0(\delta_k|\delta_1^{k-1}) = [P_{fa}]^{\delta_k}[1 - P_{fa}]^{1-\delta_k} \tag{60}$$

where as before P_{fa} (a function of the detection sequence) is obtained using (**??**).

Under H_1 (target present), the likelihood function for the kth measurement from the ith sensor becomes

$$f_1(\delta_k|\delta_1^{k-1}) = \sum_{l=0}^{1} f_1(\delta_k|u_{ik} = l)\Pr\{u_{ik} = l|\delta_1^{k-1}\} \tag{61}$$

where $\Pr\{u_{ik} = l|\delta_1^{k-1}\}$ is obtained using a Bayesian update from the prior measurements using the methodology described next.

Let $\Pr\{u_{i1}\}$ (used in (61) for $k = 1$) be given by the steady state probability of the Markov chain:

$$\Pr\{u_{i1}\} = \begin{cases} q/(p+q) & u_{i1} = 0 \\ p/(p+q) & u_{i1} = 1 \end{cases}. \tag{62}$$

As a recursive procedure Bayes' rule is applied to obtain the posterior pmf, $\Pr\{u_{ik}|\delta_1^k\}$, that will be used in (64),

$$\Pr\{u_{ik}|\delta_1^k\} = \frac{f_1(\delta_k|\delta_1^{k-1}, u_{ik})\Pr\{u_{ik}|\delta_1^{k-1})}{\sum_{l=0}^{1} f_1(\delta_1^k|\delta_1^{k-1}, u_{ik} = l)\Pr\{u_{ik} = l|\delta_1^{k-1}\}}. \tag{63}$$

The prior conditional pmf is updated for the next iteration using the Markov transition matrix,

$$\begin{pmatrix} \Pr\{u_{ik} = 0|\delta_1^{k-1}\} \\ \Pr\{u_{ik} = 1|\delta_1^{k-1}\} \end{pmatrix} = \begin{pmatrix} 1-p & q \\ p & 1-q \end{pmatrix} \begin{pmatrix} \Pr\{u_{i(k-1)} = 0|\delta_1^{k-1}\} & \Pr\{u_{i(k-1)} = 1|\delta_1^{k-1}\} \end{pmatrix} \tag{64}$$

This result is used in (61) for $k > 1$ to obtain the likelihood function under H_1.

2) Shiryaev Rule: The Shiryaev rule represents the optimal solution (under the i.i.d. assumption) to the quickest detection problem where the problem is formulated using a Bayesian approach. In this approach, there exists *a priori* information regarding the distribution of the change time.

The Shiryaev rule applies the Bayesian concept of declaring that a change in distribution has occured when the *a posteriori* probability of a change exceeds a given threshold. Assume that the *a priori* distribution of the change time k_c is given by a probability that change time is zero, and a geometric distribution of change times greater than zero:

$$\Pr\{k_c = k\} = \begin{cases} \beta_0 & k_c = 0 \\ (1-\beta_0)\rho(1-\rho)^{k-1} & k_c > 0 \end{cases} \tag{65}$$

Applying Bayes rule, the *a posteriori* probability of a change at time $k > 0$ is

$$\beta_k = \frac{[\beta_{k-1} + (1-\beta_{k-1})\rho]f_0(\delta_k|\delta_1^{k-1})}{[\beta_{k-1} + (1-\beta_{k-1})\rho]f_0(\delta_k|\delta_1^{k-1}) + (1-\beta_{k-1})(1-\rho)f_1(\delta_k|\delta_1^{k-1})} \tag{66}$$

The Shiryaev stopping rule becomes [3]

$$g_k = \ln\frac{\beta_k}{1-\beta_k} \tag{67}$$

$$= \ln(\rho + e^{g_{k-1}}) - \ln(1-\rho) + \ln\frac{f_0(\delta_k|\delta_1^{k-1})}{f_1(\delta_k|\delta_1^{k-1})} \tag{68}$$

where with each scan g_k is compared to a threshold h. If the threshold is exceeded then the track is terminated.

The use of a Bayesian test is appealing in that the *a priori* probability of a distribution change time can be related to the actual tracking problem. One can estimate β_0 based on the expected numbers of confirmed true and false tracks using the performance characteristics of the track confirmation module and the expected density of true targets in the surveillance region. Further, one can view a geometric distribution of change time as a model for the probability that a confirmed track on a true target diverges due either to a target maneuver or by being drawn off the target by incorrectly associating noise measurements to the track.

3) Model-Based Results—Track Termination: To compare the performance of the Page test, the Shiryaev test and the K/N rule, Monte Carlo simulations were performed in which detection sequences were generated with sensor measurements given by the model in [5] and using the parameters from Table II. For each hypothesis H_0 and H_1, 10^4 simulations were performed. Simulations under H_0 yield the average false track life, called average detection delay (ADD). Simulations under H_1 yield the average true track life, called average run length (ARL)[8]. Results are plotted for each track termination test over a set of threshold values. The Shiryaev test used $\beta_0 = 0.5$ and $\rho = 0.005$. Fig. 10 presents the track termination performance of each test.

The Page test performed worst over the operating range likely to be used in a track termination module of a track management system (ADD of false tracks 4–15 scan cycles). However the Page test becomes asymptotically better than the K/N rule as the Page test threshold, h, increases towards infinity, although the asymptotic behavior may not be important at the desired operating point of the system when tuned for target tracking. The Shiryaev test performed the best of the three tests considered. This is likely due to the information provided by the prior knowledge used in the Shiryaev test.

The sub-optimality of the Page test was not expected and is likely due to the HMM structure of the measurement distribution before the change. As mentioned previously, there exist no optimality proofs of the Page test when considering the specific change detection problem involving HMMs described here. Examination of the assumptions used in the optimality proofs of

[8]ADD and ARL are the standard terminology used in change detection literature

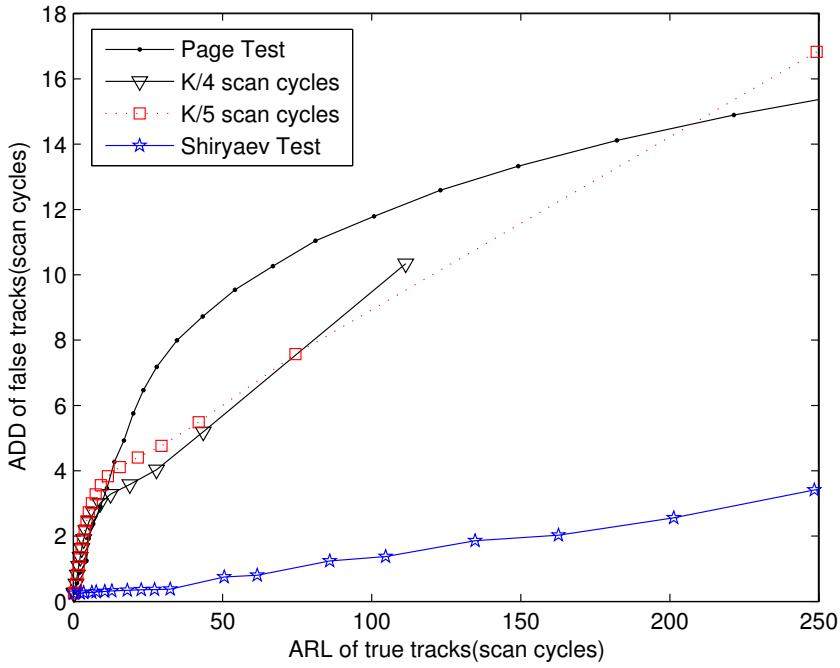

Fig. 10. Comparison of the time to terminate false and true tracks for various track termination tests.

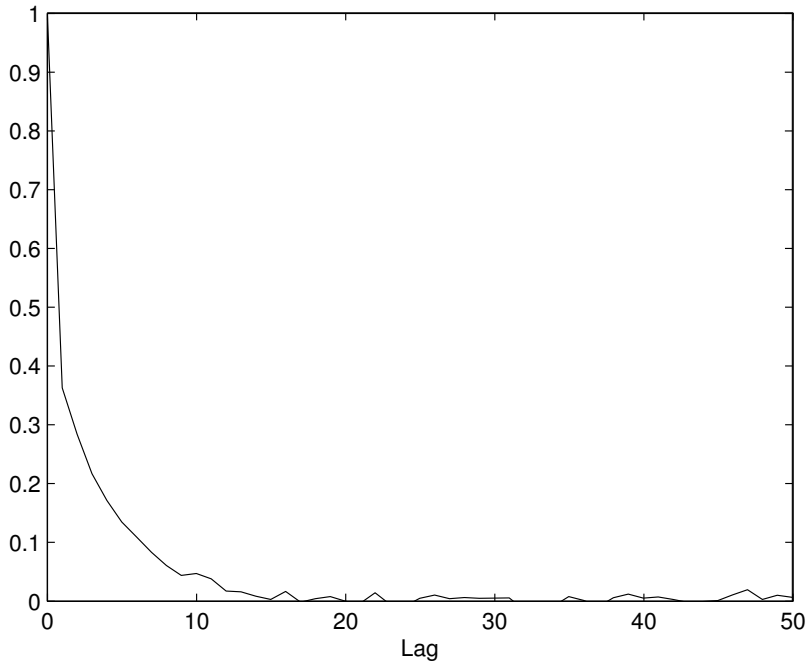

Fig. 11. Autocorrelation of the CUSUM increments, s_n, under H_1 (target present).

the Page test show that the normally claimed requirement that the measurement data be i.i.d. is overly restrictive and that optimality can be proven with the less restrictive condition that the increments of the cumulative sum, s_n, be i.i.d. [3], [22]. Fig. 11 plots the autocorrelation of the CUSUM increments as a function of lag under H_1 based on data sequences obtained through simulation and shows a significant correlation out to a lag of about 5 measurements. Therefore since the CUSUM increments are not i.i.d., existing optimality proofs of the Page test are not useful in this application. Based on this analysis, the Shiryaev test provides the best track termination performance and significantly outperforms both the K/N rule and the Page test. As a sequential test, the computational cost of the Shiryaev test is small and easily computable as part of an overall track management system.

III. Some Management Issues

A. Asynchronous vs. Synchronous Sampling for Tracking in Clutter

Many target tracking subsystems have the ability to schedule their own data rates; essentially they can "order" new information whenever they need it, and the cost is in terms of the sensor resource. But among the *un*-managed schemes, uniform sampling, in which a new measurement is requested periodically and regularly, is the most commonly-used sampling scheme; deliberately nonuniform schemes are seldom given serious consideration. Here, however, we show that such schemes may have been discarded prematurely: a nonuniform sampling can have its benefits. Two sampling schemes are compared. It will turn out that the superiority of one versus another is very much a function not only of parameters, but also of the tracking scheme used. Specifically, we shall investigate the probabilistic data association filter (PDAF) and N-D assignment algorithms [2][21][30][31]. The PDAF associates measurements to tracks one scan at a time, and hence that it will be shown that uniform sampling is always the best approach for the PDAF is perhaps unsurprising. However, when a tracker that makes use of more than one scan of past data is used (i.e. multidimensional assignment), the system with nonuniform sampling can outperform the "conventional" uniform one. The advantage becomes most obvious when the target maneuvers more, when the clutter (average number of false alarms) is high, and particularly when the probability of detection is close to unity.

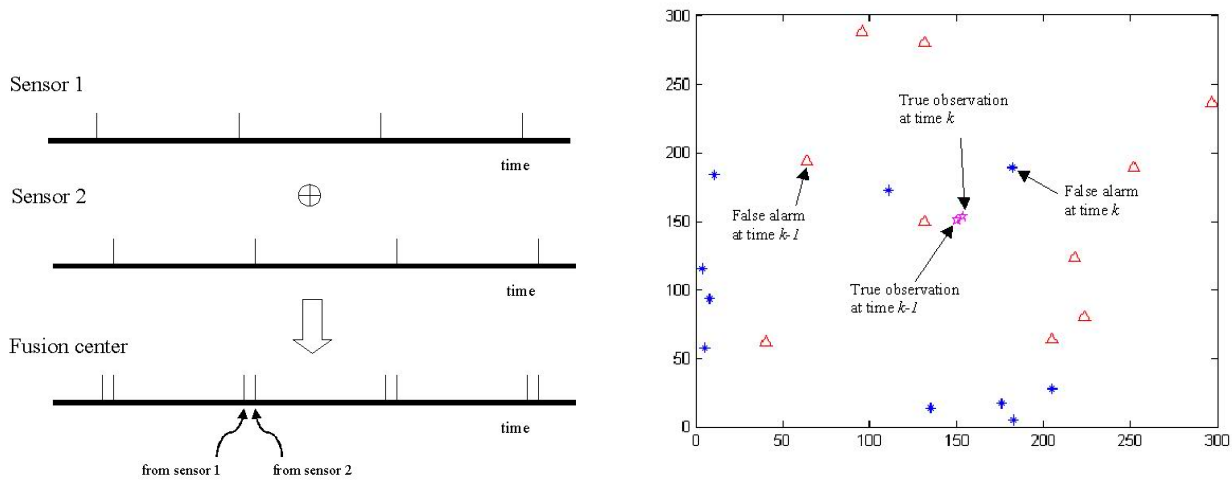

Fig. 12. Left: Possible timing of measurements from two asynchronous but equal-period sensors. Right: Overlay of true and false detections from two consecutive scans. Notice that the false measurements are unlikely to coincide. However, given that the inter-scan time is not too great, the true measurements are close, and can corroborate one another.

Consider Figure 12 (left), in which centralized processing is accomplished based on measurements from a pair of autonomous observers. These observers do not scan in lock-step, and the arrival of a measurement from either at the fusion center is aperiodic, and could even be considered (in the general case) a point-arrival random process. We do not consider explicitly the data fusion situation here: there are simply too many variables to adjust for a clear picture to emerge. Similarly, even in the single-sensor case with sensor management, we could consider more-sophisticated staggered-sampling schemes. So: is there some benefit to nonuniform sampling? The question is given detailed examination in [26], [27], [28] for multi-sensor situations without measurement-origin uncertainty. The *clean-data* assumption allows quite elegant and exact Riccati analyses: it turns out that for identical sensors a "uniform stagger" is best, although for sensors having measurements of dissimilar quality there can be benefit to a richer pattern. Here we are most interested in systems for which data association is necessary: in the presence of missed detections and false alarms, can a non-uniform sampling strategy be the best choice even for identical sensors? One reason why it might be is illustrated in Figure 12 (right): when the inter-sample interval is short, the true observations from two time-adjacent samples are close to each other. Thus the true observation can be made more distinguishable from those that are false by use of a short sampling interval: two measurements in adjacent scans that are very close are likely both true,

while any measurement in a first scan that is not repeated by one in the second scan that corroborates it is probably (although not definitely) false.

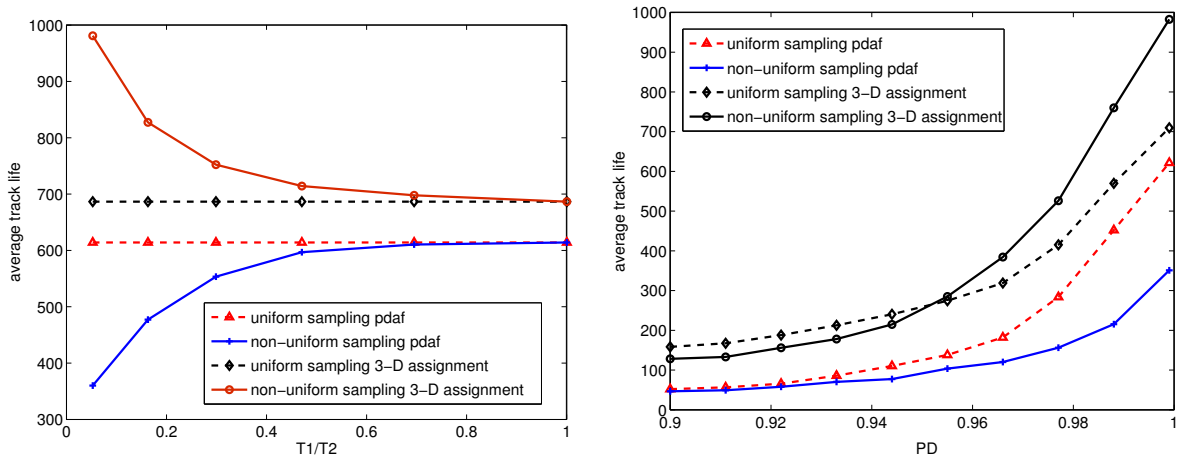

Fig. 13. Left: Average track life in PDAF and 3-D assignment, as a function of T_1/T_2, with uniform and nonuniform sampling schemes, false alarm density $\lambda = 0.01$, $T = 1s$, $P_D = 1$, $\tilde{q} = 10$, and $\sigma_w^2 = 0.1$. Right: Average track life in PDAF and 3-D assignment, as a function of probability of detection, with uniform and nonuniform sampling schemes, $T_1 = 0.1s$, $T_2 = 1.9s$, $T = 1s$, $\lambda = 0.01$, $\tilde{q} = 10$, and $\sigma_w^2 = 0.1$.

The benefit of nonuniform sampling is clear from Figure 13 (left), which shows the average track life as a function of the ratio T_1/T_2, with of course the stipulation $T_1 + T_2 = 2$ seconds in all cases; 1000 Monte Carlo runs were used to obtain the results, and details of the simulations are given in [38]. It is clear from this that when pairs of samples are close (i.e. a small ratio) there can be a significant improvement from a nonuniform approach. The difference can be substantial (fifty percent!) when the ratio approaches zero: interestingly, this is the (synchronous) "data fusion" case discussed earlier. Figure 13 (left) is particularly kind to nonuniformity because it depicts the case of perfect detection: there *are* false alarms, but no measurements are missed. Figure 13 (right) explores this further. Apparently, the improvement for the nonuniform sampling case is greatest when the target is faithfully present in the data: we shall discuss this in the next section. Generally, though, we have found that nonuniform sampling is preferable only when $P_D > 95\%$.

All in all, for a simple scan-based tracker (like the PDAF) there is no benefit from anything other than a uniform sampling rate [26], [27], [28], [38]. However, when the tracker utilizes more than one scan of data from the past — as with multiple-frame assignment or, presumably, the multi-hypothesis tracker (MHT) — nonuniform can outperform uniform sampling. This is particularly so when:

- the target is highly maneuverable; and/or
- the false alarm density is high; and/or
- the probability of detection is high.

We have measured in terms of percentage of lost tracks and average track lifetime; and in all cases the *aggregate sensor resource is kept constant*, meaning, for example, that one non-uniform scheme with short interval 0.1 second and long interval 1.9 seconds can be fairly compared to a uniform scheme in which the constant inter-sample interval is 1 second. Further, although management of the "resource" of a single sensor has provided the results here, the implications on *data fusion* should be clear. That is, given the conditions of the above three bullets, it may be better to purchase low-rate sensors that scan simultaneously and fuse, as opposed to a single high-rate sensor.

B. Multistatic Sonar Sensor Layout

In [13] we proposed an optimization technique for the optimal sensor placement for multistatic sonar systems in the LFM-only case, the CF-only case, and the combined LFM-CF case. An important aspect of the algorithm is that we employ a "minimax" criterion which results in a balanced surveillance performance. This makes sure that there is no path across the barrier for a target yet it remains "unseen". Some aspects of modeling are important:

- Targets are not "point" targets: we employ an aspect angle dependent target strength model.
- Target Doppler is included in the localization analysis whenever CF waveforms are used.
- It is assumed that targets follow some realistic trajectories; Hence, availability of two complementary waveforms, CF and LFM, is incorporated in the metric.
- The modeling reflects the "Blanking Zone" due to direct blast signal reception.
- Signal Excess is calculated by a model where a simplified reverberation-limited sonar equation is used and the Q-function helps quantify the Doppler performance of sonar waveforms in rejecting reverberation.

In finding an optimal sensor placement, the main objective was to improve target tracking performance: the "information" flow to the tracker was the basis. The Fisher information matrix can be seen as a quantification of information in the measurement about the target's state, defined as the inverse of the covariance of the estimate:

$$I(X_s, X_r, X_t, \omega) = R(X_s, X_r, X_t, \omega)^{-1} \tag{69}$$

For optimization purposes, we need a scalar quantity for each source, receiver and target configuration for a given waveform ω (CF or LFM). We use the "information gain"

$$I_{fused}(X_t) = \sum_{\forall\omega} \sum_{\forall(s,r)\in Y} P_d^{\omega}(s,r) I(X_s, X_r, X_t, \omega) \tag{70}$$

$I_{fused}(X_t)$ is a function of target location given a particular geometry Y – the locations of sources and receivers. Note that equation (70) is based on the simplifying approximation that sensor measurement errors are uncorrelated from one contact to another, and indeed can be related to the PCRLB [16] for the case of a target without process noise and in the absence of false alarms. The expression, while simple and useful for our purposes, has some degree of optimism: the true information gain is upper bounded by this expression.

Direct blast blanking means that for certain source-target-receiver geometries the detection probability that follows from our signal-excess modeling must be replaced by zero. Rather than doing so, and for numerical stability in the optimization process, we choose instead to discount the information gain with a barrier-type function. That is, as the target moves into the direct blast region, it is still detected but with a rapidly increasing localization uncertainty:

$$I(X_s, X_r, X_t, \omega) = e^{-\kappa d} \cdot I(X_s, X_r, X_t, \omega) \tag{71}$$

where d is the shortest distance between the target and the border of the blanking zone ellipse.

We choose the determinant to be the scalar measure of the quality of information available to the tracker at each waypoint. Moreover, we consider a set of linear target trajectories T, each consisting of several waypoints. The number of waypoints along each trajectory differs based on the speed of the target and the sampling interval; the latter is chosen so as to have several waypoints for the fastest-moving trajectories of interest. We use the (optimistic) simplifying approximation that information gained along a trajectory is the summation of the information across waypoints. Thus, as the scalar measure for each trajectory $T_i \in T$, we use the summation of determinants of the fused information matrix over all waypoints $w_{ij} \in T_i$:

$$M(T_i) = \sum_j det(I_{fused}(w_{ij})) \tag{72}$$

The objective function J id defined as the *worst-case* (i.e. smallest) information gain achieved across all trajectories:

$$J = \min_i M(T_i) \tag{73}$$

Maximization of the latter objective function is in fact the well-known *minimax* criterion: minimization of the maximum possible loss. In an overt network, a threat submarine would try to choose a path so that it would not be detected. Hence, operationally, the minimax criterion makes more sense since it makes sure that there are no "holes" in the surveillance region. We choose it as our objective in the optimization. Note that this objective incorporates (and maximizes) both localization accuracy and the detection opportunities over the whole trajectory of the target. In other words, it aims to improve the tracking accuracy at all instances of target penetration. Hence, it can be seen that it relates to other operationally meaningful objectives, such as maintaining (not losing) a track, or increasing the target detectability.

We refer to a source-receiver pair as a detection node and consider 2-node systems – the 3-node case is discussed in [13]. Is it better to deploy two sources with one receiver, or is the system with two-receivers and a source good enough? Where should the assets be deployed. We focus on the important "barrier" (denial of access) case, in which the barrier is the region (-35km, -10km) to (35km, 10km); there are 15 hypothetical target trajectories considered along this barrier. For instance, trajectory 1 represents a target with heading 200 degrees (from North) and $10kts$ speed. Along this trajectory, there are 5 waypoints.

Beginning with the LFM signal case (good localization but no Doppler) we consider two 2-node systems: one source and two receivers, and two sources and one receiver. It turns out that the former case is better in terms of our score function, and that in either case the optimal placement is that the sources (blue squares in the figure 14) and receivers (circled star in the figure 14) form a line in the North-South direction. This is intuitive since it allows sensors to see the target from broadside. The target strength is at its maximum if the bistatic angle is close to 90 degrees (i.e., broadside), meaning that the SNR is high. Moreover, the receivers are located so that for any given target location, the orientation of the uncertainty ellipses becomes complementary. Note also that the placement of the assets is not quite symmetric: this is a repeatable result of the optimization, and is not an accident.

Considering now CF waveforms, the optimal placements are given in figure 15. This time the sensors are in the west-east orientation. This is again intuitive since the penetrating target provides high range rate (Doppler) measurement, and hence the information provided to tracker is higher. The complementarity of the waveforms is consistent with the complementarity of the

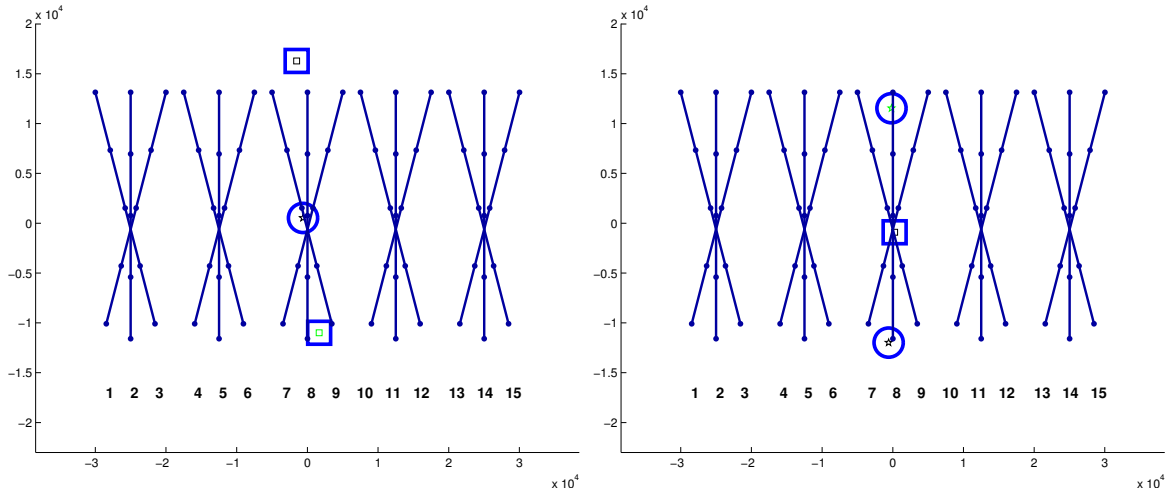

Fig. 14. The LFM case. Left: 2 Sources (the squares), 1 Receiver (the circle) case. The lines are trajectories and the dots represent waypoints of each trajectory. Targets head south. The optimal placement forms a line in the North-South direction. Right: The 1 Source, 2 Receivers case.

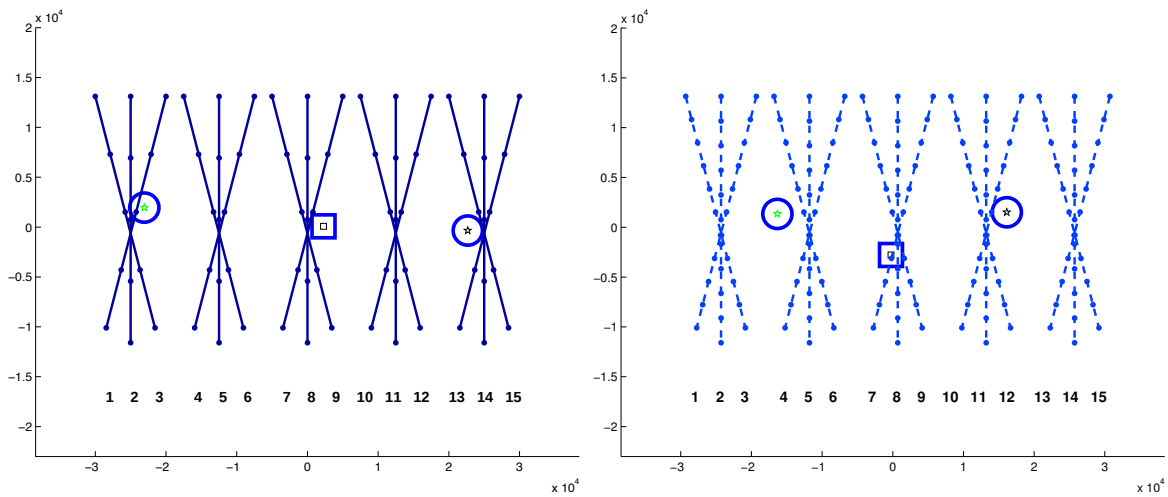

Fig. 15. The CF case, both with one source and two receivers. Left: Target speed is high, 10 knots. Right: Target speed is low, 4 knots.

optimal solutions. Another important observation is that the scores from the faster target are much better (lower) than the other one. This indicates that the Doppler information is so dominant that even though the slow target has many more waypoints and hence many more chances to be detected, it is harder to detect it if only a CF waveform is used.

C. Choice of Waveform

It is well-known that an appropriate waveform shape for Doppler sensitivity is one of constant frequency (CF); on the other hand, a CF waveforms range resolution is generally poor compared to one with higher bandwidth, and a common high-bandwidth waveform is one that uses a linearly-swept frequency-modulated (LFM) waveform. A GMTI system, for instance, is therefore faced with a trade-off: use CF to avoid clutter, and use LFM for accuracy. In Figure 16 we show a notional plot of the interaction between an LFM probe (with accuracy given by measurement error covariance $R(k)$) and kinematic model's track: the prior uncertainty ($P(k|k-1)$) of a kinematic target always has a positive correlation between range and range-rate, while the measurement uncertainty correlation between range and range-rate varies from negative (d) for LFM upsweep to zero (c) for CF to positive (b) for LFM down sweep. Clearly the LFM upsweep is best.

In [25] the impact of waveform selection and measurement extraction on radar system (i.e., tracking) performance in the case that both range and range-rate observations are available was investigated. Here are the major conclusions:

1) Because of the positive correlation between track estimation errors in position and velocity, tracking is best served by a measurement whose position and velocity errors are negatively correlated. This is easily achieved by LFM upsweep; but the "busy" nature of the ambiguity functions of many fancier waveforms (e.g. VFM, XFM, parabolic FM, and coded CF) destroys this correlation, and their effectiveness in tracking is unimpressive [6], [23]. Further, although an alternating CF/LFM waveform suite performs well, it is seldom as useful as a consistent LFM choice.

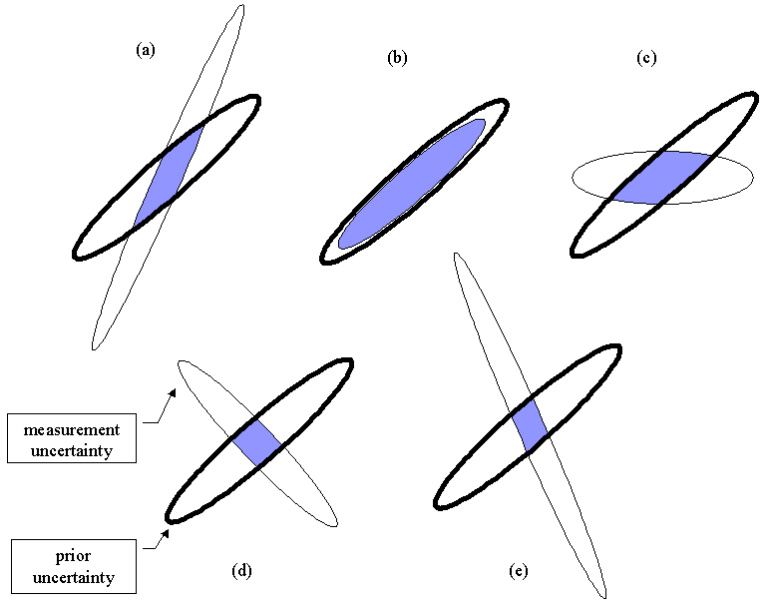

Fig. 16. Illustration of the complementarity of prior (heavy-shaded) and measurement (light-shaded) uncertainties for LFM waveforms, in range (horizontal) and range-rate (vertical) coordinates. The blue shaded region is a cartoon of the posterior uncertainty.

2) For each target kinematic model, as expressed by its maneuvering index, there exists a tracking-optimal LFM upsweep frequency rate. This frequency sweep rate increases as the target motion model becomes closer to constant velocity (lower maneuvering index).
3) It is important to control ambiguity function sidelobes. There can be as much as a factor of two improvement when Hamming weighting is used, as compared with the rectangular-envelope case. Hamming weighting appears to be a good choice, in general.
4) The probability of detection for a target not precisely at the point of sampling (i.e. not at the center of a resolution cell) is degraded. Thus, it is worth considering the use of a synthetic discriminant function (SDF) mismatched filter, which is robust to such perturbations. Results indicate that the performance is SNR-dependent: at high SNR it offers a distinct improvement, while at low return strength the ordinary matched filter is better.
5) We have in past work used a measurement extraction system by which contiguous threshold-exceedances (in neighboring resolution cells) are combined to a single measurement, presumably of enhanced accuracy. We have compared this to a *strongest-neighbor* approach, and to an amplitude-weighted average; the direct average is a significant improvement over the strongest-neighbor scheme. The amplitude-weighted is essentially identical in performance to the direct-average procedure for low-observable targets, but there is a marked improvement in the use of the more-involved amplitude-weighted approach as the SNR grows larger than 20dB.
6) A simple alternative to these measurement-extraction approaches is direct use of the data association algorithm, basically to give the PDAF target tracker all threshold exceedances and to allow it to extract a combined measurement. Our results show that there is very little loss from the use of this approach – we consider this surprising as the PDAF is designed predicated on an assumption of at most one measurement per target. We note that most alternative tracking algorithms (MHT and assignment) do not immediately allow for "soft" association of multiple returns, and hence in these cases some form of measurement extraction is necessary.

Finally, we note that a popular scheme involves the use of both upsweep and downsweep LFM waveforms — from [14] it was clear that by simple averaging the effect of the target's range-rate can be cancelled. This is the VFM discussed in [6], and when both range and range-rate measurements are available VFM does not perform as well as a consistent use of LFM upsweep. In fact, we conclude that correlated range and range-rate measurements are desirable and should not be avoided.

REFERENCES

[1] Y. Bar-Shalom, X. Li, and T. Kirubarajan, *Estimation with Application to Tracking and Navigation*, Wiley, 2001.

[2] Y. Bar-Shalom and X. Li, *Multitarget-Multisensor Tracking: Principles and Techniques*, YBS Publishing, 1995.

[3] M. Basseville and L. Nikiforov, *Detection of Abrupt Changes*. New Jersey: Prentice Hall, 1993

[4] S. Blackman and R. Popoli, *Design and Analysis of Modern Tracking Systems*, Artech House (Boston), 1999.

[5] W. Blanding, P. Willett, Y. Bar-Shalom and S. Coraluppi, "Multisensor Track Management for Targets with Fluctuating SNR," *to appear in IEEE Transactions on Aerospace and Electronic Systems*.

[6] W. Cao, P. Willett, Y. Bar-Shalom, and R. Niu, "Waveform Selection Based on Detection-Tracking Performance", *Proceedings of the SPIE Conference on Signal and Data Processing of Small Targets*, Orlando FL, April 1998.

[7] B. Chen and P. Willett, "Quickest detection of hidden markov models," *Proc. of the 36th IEEE Conference on Decision and Control*, San Diego, CA, Dec. 1997.

[8] S. Coraluppi and C. Carthel, Distributed Tracking in Multistatic Sonar, *IEEE Trans. Aerosp. Electronic Systems*, vol. 41, pp. 11381147, July 2005.

[9] S. Coraluppi, "Localization and fusion in multistatic sonar," *Proc. of the 8th International Conference on Information Fusion*, Philadelphia, PA, July 2005.

[10] P. Davis and P. Rabinowitz, *Methods of Numerical Integration*, Second Edition, Academic Press, 1984.

[11] F. E. Daum, "Bounds on the Performance of Multiple Target Tracking", *IEEE Transactions on Automatic Control*, Vol. 35, No. 4, pp. 443-446, April 1990.

[12] F. E. Daum, "A Cramér-Rao Bound for Multiple Target Tracking", *Signal and Data Processing of Small Targets*, SPIE Proc. 1481, pp. 341-344, Orlando, FL, April 1991.

[13] O. Erdinc, P. Willett and S. Coraluppi, "Multistatic Sensor Placement: A Tracking Approach," *Journal of Advances in Information Fusion*, Vol. 2, Number 1, pp. 22-34, July 2007.

[14] Fitzgerald, R., "Effects of Range-Doppler Coupling on Chirp Radar Tracking Accuracy," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-10, pp. 528–532, July 1974.

[15] C-D. Fuh, "SPRT and CUSUM in hidden markov models," *The Annals of Statistics*, vol. 31, pp. 942-977, June 2003.

[16] M. Hernandez, P. Horridge
Advances in the management of multi-sensor systmes with associated applications
Target tracking 2004: Algorithms and allocations, IEE, 23-24 May 2004.

[17] C. Jauffret and Y. Bar-Shalom, "Track Formation with Bearing and Frequency Measurements in Clutter", *IEEE Transactions on Aerospace and Electronic Systems*, vol. 26, No. 6, pp. 999–1009, November 1990.

[18] T. Kirubarajan and Y. Bar-Shalom, "Low Observable Target Motion Analysis Using Amplitude Information", *IEEE Transactions on Aerospace and Electronic Systems*, vol. 32, No. 4, pp. 1367–1384, October 1996.

[19] X. R. Li and Y. Bar-Shalom, "A Hybrid Technique for Performance Evaluation of Tracking in Clutter", *IEEE Transactions on Automatic Control*, vol. AC-36, pp. 588–602, May 1991.

[20] L. Ljung, *System Identification: Theory for the User*, Prentice-Hall, 1987.

[21] C.L. Morefield, "Application of 0-1 Integer Programming to Multi-target Tracking Problems", *IEEE Transactions on Automatic Control*, Vol. 22, pp. 302-312, June 1977.

[22] G. Moustakides, "Quickest detection of abrupt changes for a class of random processes," *IEEE Trans. Inform. Theory*, vol. 44, pp. 1965-1968, Sept. 1998.

[23] R. Niu, P. Willett, and Y. Bar-Shalom, "From the Waveform through the Resolution Cell to the Tracker", *Proceedings of the 1999 Aerospace Conference*, Snowmass CO, March 1999.

[24] R. Niu, P. Willett, and Y. Bar-Shalom, "Matrix CRLB Scaling Due to Measurements of Uncertain Origin", *IEEE Transactions on Signal Processing*, vol. 49, No. 7, pp. 1325–1335, July 2001.

[25] R. Niu, P. Willett and Y. Bar-Shalom, "Tracking Considerations in Selection of Radar Waveform Given Range and Range-Rate Measurements," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 38, No. 2, pp. 467-487, April 2002.

[26] R. Niu, P. Varshney, K. Mehrotra and C. Mohan, "Temporal Fusion in Multi-Sensor Target Tracking Systems," *Proceedings of the International Conference on Information Fusion*, Annapolis MD, July 2002.

[27] R. Niu, P. Varshney, K. Mehrotra and C. Mohan, "Sensor Staggering in Multi-Sensor Target Tracking Systems," *Proceedings of the 2003 IEEE Radar Conference*, Huntsville AL, May 2003.

[28] R. Niu, P. Varshney, K. Mehrotra and C. Mohan, "On Temporally Staggered Sensors in Multi-Sensor Target Tracking Systems," *IEEE Transactions on Aerospace and Electronic Systems*, October 2005.

[29] K. R. Pattipati, R. L. Popp and T. Kirubarajan, "Survey of Assignment Techniques for Multitarget Tracking", Chapter 4 of *Multitarget-Multisensor Tracking: Applications and Advances, Volume III, Yaakov Bar-Shalom and William Dale Blair Editors*, Artech House, 2000.

[30] K.R. Pattipati, S. Deb, Y. Bar-Shalom, and R.B. Washburn, "Passive Multisensor Data Association Using a New Relaxation Algorithm", In *Multitarget-Multisensor Tracking: Advanced Applications*, Y. Bar-Shalom, editor. Artech House, 1990.

[31] K.R. Pattipati, S. Deb, Y. Bar-Shalom, and R.B. Washburn, "A New Relaxation Algorithm and Passive Sensor Data Association", *IEEE Transactions on Automatic Control*, Vol. 37(2), February 1992.

[32] Poor, H., *An Introduction to Signal Detection and Estimation*, 2nd ed. New York: Springer, 1994.

[33] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, *Numerical Recipies in C*, Cambridge, 1992.

[34] M. Smith and E. Winter, "On the detection of target trajectories in a multi-target environment," *Proc. of the IEEE Conference on Decision and Control*, San Diego, CA, 1979, pp. 11891194.

[35] P. Tichavský, C. Muravchik, A. Nehorai, "Posterior Cramér-Rao Bounds for Discrete-Time Nonlinear Filtering", *IEEE Trans. Signal Processing*, vol. 46, No. 5, pp. 1386–1396, May 1998.

[36] H. Van Trees, *Detection, Estimation, and Modulation Theory, Part I*, Wiley, 1968.

[37] X. Zhang, P. Willett and Y. Bar-Shalom, "A Dynamic Cramer-Rao Lower Bound for Target Tracking in Clutter," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 41, No. 4, pp. 1154-1157, October 2005.

[38] X. Zhang, P. Willett and Y. Bar-Shalom, "Uniform Versus Nonuniform Sampling When Tracking in Clutter," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 42, No. 2, pp. 388-400, April 2006.